

DECOMPOSITION OF GEODESICS IN THE WASSERSTEIN SPACE AND THE GLOBALIZATION PROPERTY

FABIO CAVALLETTI

ABSTRACT. Let (X, d, m) be a non-branching metric measure space verifying $\text{CD}_{\text{loc}}(K, N)$ or equivalently $\text{CD}^*(K, N)$. In this note we show that given a geodesic μ_t in the L^2 -Wasserstein space, it is always possible to write the density of μ_t as the product of two densities, one corresponding to a geodesic with support of codimension one verifying $\text{CD}(K, N - 1)$, and the other associated with a one dimensional measure.

For a particular class of optimal transportation we prove the linearity in time of the other component, obtaining therefore the full $\text{CD}(K, N)$ for μ_t . This result can be therefore interpret as the “self-improving property” for $\text{CD}^*(K, N)$ or as a partial globalization theorem for $\text{CD}(K, N)$.

In the setting of infinitesimally strictly convex metric measure space, we also write explicitly the one dimensional density obtaining a complete and explicit decomposition of the density.

1. INTRODUCTION

An important class of singular spaces is the one of metric measure spaces with generalized lower bounds on the Ricci curvature formulated in terms of optimal transportation. This class of spaces, together with lower bounds on curvature and upper bounds on dimension, has been introduced by Sturm in [13, 14] and independently by Lott and Villani in [12] and it is called $\text{CD}(K, N)$.

The *curvature-dimension condition* $\text{CD}(K, N)$ depends on two parameters K and N , playing the role of a curvature and dimension bound, respectively. We recall two important properties of the condition $\text{CD}(K, N)$:

- the curvature-dimension condition is stable under convergence of metric measure spaces with respect to the L^2 -transportation distance \mathbb{D} introduced in [13];
- a complete Riemannian manifold satisfies $\text{CD}(K, N)$ if and only if its Ricci curvature is bounded from below by K and its dimension from above by N .

Moreover a broad variety of geometric and functional analytic properties can be deduced from the curvature-dimension condition $\text{CD}(K, N)$: the Brunn-Mikowski inequality, the Bishop-Gromov volume comparison theorem, the Bonnet-Myers theorem, the doubling property and local Poincaré inequalities on balls. All this listed results are quantitative results (volume of intermediate points, volume growth, upper bound on the diameter and so on) depending on K, N .

Curvature-dimension condition $\text{CD}(K, N)$ prescribes how the volume of a given set is affected by curvature when it is moved via optimal transportation. Condition $\text{CD}(K, N)$ impose that the distortion is ruled by the coefficient $\tau_{K,N}^{(t)}(\theta)$ depending on the curvature K , on the dimension N , on the time of the evolution t and on the point length θ . The main feature of $\tau_{K,N}^{(t)}(\theta)$ is that it is obtained mixing two different volume distortions: an $(N - 1)$ -dimensional distortion depending on the curvature K and a one dimensional evolution that doesn't feel the curvature. Namely

$$\tau_{K,N}^{(t)}(\theta) = t^{1/N} \sigma_{K,N-1}^{(t)}(\theta)^{(N-1)/N},$$

where $\sigma_{K,N-1}^{(t)}(\theta)^{(N-1)/N}$ contains the information on the $(N - 1)$ -dimensional volume distortion and the evolution in the remaining direction is ruled just by $t^{1/N}$. That was inspired by the Riemannian framework. The coefficient $\sigma_{K,N-1}^{(t)}(\theta)$ are the fundamental solutions of a second order ordinary differential equation with parameters K and N .

One of the most important questions on $\text{CD}(K, N)$ that are still open, is whether this notion enjoys a globalization property: can we say that a metric measure space (X, d, m) satisfies $\text{CD}(K, N)$ provided $\text{CD}(K, N)$ holds true locally on a family of sets X_i covering X ?

A first tentative of answer this problem was given by Bacher and Sturm in [4]: they proved that a metric measure space (X, d, m) verifies the local curvature-dimension condition $\text{CD}_{loc}(K, N)$ if and only if it verifies the global reduced curvature-dimension condition $\text{CD}^*(K, N)$. The latter is obtained from $\text{CD}(K, N)$ imposing that the volume distortion, during the evolution through an optimal transportation, is ruled by $\sigma_{K,N}^{(t)}(\theta)$ instead of $\tau_{K,N}^{(t)}(\theta)$. The reason why this result holds true stays in the better behavior (in time) of $\sigma_{K,N}^{(t)}(\theta)$ than $\tau_{K,N}^{(t)}(\theta)$.

The reduced condition is strictly weaker than $\text{CD}(K, N)$ and a converse implication can be obtained only changing the value of the lower bound on the curvature: condition $\text{CD}^*(K, N)$ implies $\text{CD}(K^*, N)$ where $K^* = K(N - 1)/N$. Therefore $\text{CD}^*(K, N)$ gives worse geometric and analytic information than $\text{CD}(K, N)$.

In [8] the author and Sturm proved that if (X, d, m) is a non-branching metric measure space that verifies $\text{CD}_{loc}(K, N)$ then (X, d, m) verifies $\text{MCP}(K, N)$, where the latter is a weaker variant of $\text{CD}(K, N)$: while $\text{CD}(K, N)$ is a condition on the optimal transport between any pair of absolutely continuous (w.r.t. m) probability measure on X , $\text{MCP}(K, N)$ is a condition on the optimal transport between a Dirac delta and the uniform distribution m on X .

The approach in [8] was to isolate a local $(N - 1)$ -dimensional evolution ruled by $\sigma_{K,N-1}^{(t)}(\theta)$ and then using the nice properties of $\sigma_{K,N-1}^{(t)}(\theta)$, obtain a global $(N - 1)$ -dimensional evolution with coefficient $\sigma_{K,N-1}^{(t)}(\theta)$. Then using Hölder inequality and the linear behavior of the other direction, pass from the $(N - 1)$ -dimensional version to the full-dimensional version with coefficient $\tau_{K,N}^{(t)}(\theta)$.

However to detect the $(N - 1)$ -dimensional evolution it is necessary to decompose the whole evolution. Considering the optimal transport between a Dirac mass in o and the uniform distribution m , the family of spheres around o immediately provides the correct $(N - 1)$ -dimensional evolutions.

The aim of this paper is to identify, in the general case of optimal transportation between any measures, the $(N - 1)$ -dimensional evolution verifying $\text{CD}^*(K, N - 1)$ and, starting from that, provide a decomposition for densities of geodesics. The density will be written as the product between the $(N - 1)$ -dimensional density verifying $\text{CD}^*(K, N - 1)$ and of a 1-dimensional density not necessarily associated to a 1-dimensional geodesic. In the framework of infinitesimally strictly convex spaces, the 1-dimensional density will be explicitly written producing an explicit decomposition.

Applications to the globalization problem for $\text{CD}(K, N)$ are also given. With this approach we are able to reduce the problem to prove concavity in time of the 1-dimensional density. In the particular case of optimal transport plans giving the same speed to geodesics leaving from the same level set of the associated Kantorovich potential, we prove indeed linearity of the 1-dimensional factor and we get the full $\text{CD}(K, N)$ inequality.

We now present the results contained in this note in more details. The first difficulty we have to handle with is to find a suitable partition of the space. Unlikely optimal transportations connecting measures to deltas, there is not just a universal family of sets but one for each $t \in [0, 1]$: if μ_t is a geodesic in W_2 and φ the Kantorovich potential associated, then

$$\{\gamma_t : \varphi(\gamma_0) = a, \gamma \in \text{supp}(\gamma)\}_{a \in \mathbb{R}}$$

is the family of partitions, one for each $t \in [0, 1]$, that will be considered. Here $\gamma \in \mathcal{P}(\mathcal{G}(X))$ is the dynamical optimal transference plan of μ_t . The reason why that family is the right one stays in the following property: the set

$$\{(\gamma_0, \gamma_1) \in X \times X : \varphi(\gamma_0) = a\}$$

is d -cyclically monotone (Lemma 4.1). Hence for $\gamma \neq \hat{\gamma} \in \text{supp}(\gamma)$ with $\varphi(\gamma_0) = \varphi(\gamma_1)$ it holds

$$\gamma_s \neq \gamma_t, \quad \forall s, t \in (0, 1),$$

therefore for $s \neq t$, $\{\gamma_s : \varphi(\gamma_0) = a\}$ and $\{\gamma_t : \varphi(\gamma_0) = a\}$ are disjoint. This key property permits to consider the evolution of the geodesic when restricted on the level sets of Kantorovich potential.

To perform a dimensional reduction argument on measures the right tool is Disintegration Theorem 2.18: if $\gamma \in \mathcal{P}(\mathcal{G}(X))$ is the dynamical optimal transference plan, then (Proposition 3.1)

$$\gamma = \int_{\varphi(\mu_0)} \gamma_a \mathcal{L}^1(da), \quad \gamma_a(\{\gamma \in G : \varphi(\gamma_0) = a\}) = \|\gamma_a\|,$$

where $\varphi(\mu_0) = \varphi(\text{supp}(\mu_0))$ and G is the support of γ . Clearly the geodesic to consider is $t \mapsto (e_t)_\#(\gamma_a)$ but, being singular with respect to m , it is not clear which reference measures we have to choose. One option could be to consider for each $t \in [0, 1]$, the family

$$\{\gamma_t : \varphi(\gamma_0) = a, \gamma \in G\}_{a \in \varphi(\mu_0)}.$$

For each $t \in [0, 1]$, by d^2 -cyclical monotonicity, the family is a partition of $e_t(G)$ and hence we have (Proposition 3.1 and Lemma 3.2)

$$m_{\perp e_t(G)} = \int_{\varphi(\mu_0)} \hat{m}_{a,t} \mathcal{L}^1(da), \quad \hat{m}_{a,t}(\{\gamma_t : \varphi(\gamma_0) = a\}) = \|\hat{m}_{a,t}\|.$$

But if we want to use the $(N-1)$ -dimensional measures $\hat{m}_{a,t}$ as reference measures, there is then no way to prove a $CD^*(K, N-1)$ estimate for the densities of γ_a , indeed if $(e_t)_\# \gamma = \mu_t = \varrho_t m$, then, up to renormalization constant,

$$\gamma_a = \varrho_t \hat{m}_{a,t},$$

and therefore the density is ϱ_t .

The correct reference measures are built as follows. For each $a \in \varphi(\mu_0)$, consider the following family of sets

$$\{\gamma_t : \varphi(\gamma_0) = a, \gamma \in G\}_{t \in [0, 1]}.$$

By d -cyclical monotonicity, they are disjoint (Lemma 4.2). If $\bar{\Gamma}_a(1) := \cup_{t \in [0, 1]} \{\gamma_t : \varphi(\gamma_0) = a, \gamma \in G\}$, then (Proposition 4.3)

$$m_{\perp \bar{\Gamma}_a(1)} = \int_{[0, 1]} m_{a,t} \mathcal{L}^1(dt), \quad m_{a,t}(\{\gamma_t : \varphi(\gamma_0) = a\}) = \|m_{a,t}\|.$$

For every $t \in [0, 1]$, we have $(e_t)_\#(\gamma_a) \ll m_{a,t}$ (Corollary 5.6). If $h_{a,t}$ is the corresponding density, then $t \mapsto h_{a,t}(\gamma_t)$ satisfies the local reduced curvature-dimension condition $\text{CD}_{loc}^*(K, N-1)$ (Theorem 6.4). Here the main idea is to consider a geodesic in the Wasserstein space, absolute continuous with respect to m , moving in the same direction of $t \mapsto (e_t)_\# \gamma_a$. Taking inspiration from the Riemannian framework, the volume distortion affects only $(N-1)$ dimensions.

The geodesic is built as follows: let $0 < R_0, L_0, R_1, L_1 < 1$ with $R_0 < R_1$ be real numbers such that $R_t + L_t < 1$ for all $t \in [0, 1]$, where $R_t := (1-t)R_0 + tR_1$ and $L_t := (1-t)L_0 + tL_1$. Then

$$t \mapsto \nu_t = \frac{1}{L_t} \int_{(R_t, R_t+L_t)} (e_s)_\# \gamma_a \mathcal{L}^1(ds) \in \mathcal{P}(M)$$

is a geodesic (Lemma 6.1 and Lemma 6.2). Suitably choosing L_0 and L_1 , we prove that $t \mapsto h_{a,t}(\gamma_t)$ satisfies the local reduced curvature-dimension condition $\text{CD}_{loc}^*(K, N-1)$. It is then fairly easy to pass to the global $\text{CD}^*(K, N-1)$ (Theorem 7.2). Lemma 6.1, Lemma 6.2 and Theorem 7.2 were already presented in [8].

So up to normalization constant

$$(e_t)_\#(\gamma_a) = \varrho_t \hat{m}_{a,t} = h_{a,t} m_{a,t},$$

with $h_{a,t}$ verifying $\text{CD}^*(K, N-1)$. Therefore in order to have a complete splitting of ϱ_t in two densities, it is necessary to write explicitly $m_{a,t}$ in terms of $\hat{m}_{a,t}$ or vice versa. Note that both measures are concentrated on the set

$$\{\gamma_t : \varphi(\gamma_0) = a, \gamma \in G\}.$$

Through a careful blow-up analysis (Proposition 5.2 and Theorem 5.5), we show that the two measures are equivalent, $\hat{m}_{a,t} = \lambda_t m_{a,t}$, and we give also an explicit formula for the density λ_t (Corollary 5.4). Define $\Phi_t : e_t(G) \rightarrow \mathbb{R}$ as follows: $\Phi_t(\gamma_t) = \varphi(\gamma_0) + \frac{t}{2} d^2(\gamma_0, \gamma_1)$, then

$$\frac{1}{\lambda_t(\gamma_t)} = \liminf_{\tau \rightarrow 0} \frac{\Phi_t(\gamma_t) - \Phi_t(\gamma_{t+\tau})}{\tau}.$$

So we have proven the following decomposition of ϱ_t :

$$\varrho_t(\gamma_t) = \left(\int \varrho_t(z) \hat{m}_{a,t}(dz) \right) \frac{1}{\lambda_t(\gamma_t)} h_{a,t}(\gamma),$$

where $a = \varphi(\gamma_0)$. Note that the value of the integral does not depend on time, but just on a , hence, in order to prove $\text{CD}(K, N)$ -like estimates, the integral can be dropped out.

The last part of this note is devoted to the study of λ_t . Assuming that $d(\gamma_0, \gamma_1)$ is a function of $\varphi(\gamma_0)$, meaning that γ gives the same speed to geodesics starting on the same level set of φ , we are able to prove (Proposition 8.1) that

$$\lambda_t(\gamma_t) = (1 - t)\lambda_0(\gamma_0) + t\lambda_1(\gamma_1).$$

It is then fairly easy to get under the same assumption the global $\text{CD}(K, N)$ (Theorem 8.4). Without doing any assumption on the transference plan but assuming the space to be infinitesimally strictly convex (see (2.7)), we prove that (Proposition 8.3)

$$\frac{1}{\lambda_t(\gamma_t)} = -D\Phi_t(\nabla\varphi_t)(\gamma_t), \quad \gamma - a.e.\gamma,$$

and hence the general decomposition: up to a constant (in time) factor

$$\varrho_t = -D\Phi_t(\nabla\varphi_t)h_t.$$

Even if the starting hypothesis of this note can be chosen to be equivalently $\text{CD}_{loc}(K, N)$ or $\text{CD}^*(K, N)$, the results proved can be read from two different prospective, accordingly to $\text{CD}_{loc}(K, N)$ or $\text{CD}^*(K, N)$. From the point of view of $\text{CD}^*(K, N)$, where the globalization property is already known, the main result is that for nice optimal transportations the entropy inequality can be improved to the curvature-dimension condition, giving a “self-improving” type of result (see [5]). From the point of view of $\text{CD}_{loc}(K, N)$ clearly the main issue is the globalization problem. Here the main statement is that the local-to-global property is true for nice optimal transportations and the statement in the general case is strongly linked to the concavity of the 1-dimensional density λ_t . The latter it is in turn strongly linked to the composition property of the differential operator D .

2. PRELIMINARIES

Let (X, d) be a metric space. The length $L(\gamma)$ of a continuous curve $\gamma : [0, 1] \rightarrow X$ is defined as

$$L(\gamma) := \sup \sum_{k=1}^n d(\gamma(t_{k-1}), \gamma(t_k))$$

where the supremum runs over $n \in \mathbb{N}$ and over all partitions $0 = t_0 < t_1 < \dots < t_n = 1$. Note that $L(\gamma) \geq d(\gamma(0), \gamma(1))$. A curve is called *geodesic* if and only if $L(\gamma) = d(\gamma(0), \gamma(1))$. If this is the case, we can assume γ to have constant speed, i.e. $L(\gamma|_{[s,t]}) = |s - t|L(\gamma) = |s - t|d(\gamma(0), \gamma(1))$ for every $0 \leq s \leq t \leq 1$.

Denote by $\mathcal{G}(X)$ the space of geodesic $\gamma : [0, 1] \rightarrow X$ in X , regarded as subset of $C([0, 1], M)$ of continuous functions equipped with the topology of uniform convergence.

(X, d) is said to be a *length space* if and only if for every $x, y \in X$,

$$d(x, y) = \inf L(\gamma)$$

where the infimum runs over all continuous curves joining x and y . It is said to be a *geodesic space* if all x and y are connected by a geodesic. A point z will be called *t-intermediate point* of points x and y if $d(x, z) = td(x, y)$ and $d(z, y) = (1 - t)d(x, y)$.

Definition 2.1. A geodesic space (X, d) is *non-branching* if and only if for every $r \geq 0$ and $x, y \in X$ such that $d(x, y) = r/2$, the set

$$\{z \in X : d(x, z) = r\} \cap \{z \in X : d(y, z) = r/2\}$$

consists of a single point.

Throughout the following we will denote by $B_r(z)$ the open ball of radius r centered in z . A standard map in optimal transportation is the evaluation map: for a fixed $t \in [0, 1]$, $e_t : \mathcal{G}(X) \rightarrow X$ is defined by $e_t(\gamma) := \gamma_t$. The push-forward of a given measure, say η , via a map f will be denoted by $f_\# \eta$ and is defined by $f_\# \eta(A) := \eta(f^{-1}(A))$, for any measurable A .

2.1. Geometry of metric measure spaces. What follows is contained [14].

A *metric measure space* is a triple (X, d, m) where (X, d) is a complete separable metric space and m is a locally finite measure (i.e. $m(B_r(x)) < \infty$ for all $x \in X$ and all sufficiently small $r > 0$) on X equipped with its Borel σ -algebra. We exclude the case $m(X) = 0$. A *non-branching* metric measure space will be a metric measure space (X, d, m) such that (X, d) is a non-branching geodesic space.

$\mathcal{P}_2(X, d)$ denotes the L^2 -Wasserstein space of Borel probability measures on X and d_W the corresponding L^2 -Wasserstein distance. The subspace of m -absolutely continuous measures is denoted by $\mathcal{P}_2(X, d, m)$.

The following are well-known results in optimal transportation theory and are valid for general metric measure spaces.

Lemma 2.2. *Let (X, d, m) be a metric measure space. For each geodesic $\Gamma : [0, 1] \rightarrow \mathcal{P}_2(X, d)$ there exists a probability measure γ on $\mathcal{G}(X)$ such that*

- $e_{t\#}\gamma = \Gamma(t)$ for all $t \in [0, 1]$;
- for each pair (s, t) the transference plan $(e_s, e_t)_\# \gamma$ is an optimal coupling for d_W .

Consider the Rényi entropy functional

$$\mathcal{S}_N(\cdot | m) : \mathcal{P}_2(X, d) \rightarrow \mathbb{R}$$

with respect to m , defined by

$$(2.1) \quad \mathcal{S}_N(\mu | m) := - \int_X \varrho^{-1/N}(x) \mu(dx)$$

for $\mu \in \mathcal{P}_2(X)$, where ϱ is the density of the absolutely continuous part μ^c in the Lebesgue decomposition $\mu = \mu^c + \mu^s = \varrho m + \mu^s$.

Given two numbers $K, N \in \mathbb{R}$ with $N \geq 1$, we put for $(t, \theta) \in [0, 1] \times \mathbb{R}_+$,

$$(2.2) \quad \tau_{K,N}^{(t)}(\theta) := \begin{cases} \infty, & \text{if } K\theta^2 \geq (N-1)\pi^2, \\ t^{1/N} \left(\frac{\sin(t\theta\sqrt{K/(N-1)})}{\sin(\theta\sqrt{K/(N-1)})} \right)^{1-1/N} & \text{if } 0 < K\theta^2 \leq (N-1)\pi^2, \\ t & \text{if } K\theta^2 < 0 \text{ or} \\ & \text{if } K\theta^2 = 0 \text{ and } N = 1, \\ t^{1/N} \left(\frac{\sinh(t\theta\sqrt{-K/(N-1)})}{\sinh(\theta\sqrt{-K/(N-1)})} \right)^{1-1/N} & \text{if } K\theta^2 \leq 0 \text{ and } N > 1. \end{cases}$$

That is, $\tau_{K,N}^{(t)}(\theta) := t^{1/N} \sigma_{K,N-1}^{(t)}(\theta)^{(N-1)/N}$ where

$$\sigma_{K,N}^{(t)}(\theta) = \frac{\sin(t\theta\sqrt{K/N})}{\sin(\theta\sqrt{K/N})},$$

if $0 < K\theta^2 < N\pi^2$ and with appropriate interpretation otherwise. Moreover we put

$$\varsigma_{K,N}^{(t)}(\theta) := \tau_{K,N}^{(t)}(\theta)^N.$$

The coefficients $\tau_{K,N}^{(t)}(\theta)$, $\sigma_{K,N}^{(t)}(\theta)$ and $\varsigma_{K,N}^{(t)}(\theta)$ are the volume distortion coefficients with K playing the role of curvature and N the one of dimension.

The curvature-dimension condition $\text{CD}(K, N)$ is defined in terms of convexity properties of the entropy functional. In the following definitions K and N will be real numbers with $N \geq 1$.

Definition 2.3 (Curvature-Dimension condition). We say that (X, d, m) satisfies $\text{CD}(K, N)$ if and only if for each pair $\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)$ there exists an optimal coupling π of $\mu_0 = \varrho_0 m$ and $\mu_1 = \varrho_1 m$, and a geodesic $\Gamma : [0, 1] \rightarrow \mathcal{P}_2(X, d, m)$ connecting μ_0 and μ_1 such that

$$(2.3) \quad \begin{aligned} \mathcal{S}_{N'}(\Gamma(t) | m) \leq & - \int_{X \times X} \left[\tau_{K,N'}^{(1-t)}(d(x_0, x_1)) \varrho_0^{-1/N'}(x_0) \right. \\ & \left. + \tau_{K,N'}^{(t)}(d(x_0, x_1)) \varrho_1^{-1/N'}(x_1) \right] \pi(dx_0 dx_1), \end{aligned}$$

for all $t \in [0, 1]$ and all $N' \geq N$.

The following is a variant of $\text{CD}(K, N)$ and it has been introduced in [4].

Definition 2.4 (Reduced Curvature-Dimension condition). We say that (X, d, m) satisfies $\text{CD}^*(K, N)$ if and only if for each pair $\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)$ there exists an optimal coupling π of $\mu_0 = \varrho_0 m$ and $\mu_1 = \varrho_1 m$, and a geodesic $\Gamma : [0, 1] \rightarrow \mathcal{P}_2(X, d, m)$ connecting μ_0 and μ_1 such that (2.3) holds true for all $t \in [0, 1]$ and all $N' \geq N$ with the coefficients $\tau_{K, N}^{(t)}(d(x_0, x_1))$ and $\tau_{K, N}^{(1-t)}(d(x_0, x_1))$ replaced by $\sigma_{K, N}^{(t)}(d(x_0, x_1))$ and $\sigma_{K, N}^{(1-t)}(d(x_0, x_1))$, respectively.

For both definitions there is a local version. Here we state only the local counterpart of $\text{CD}(K, N)$, being clear what would be the one for $\text{CD}^*(K, N)$.

Definition 2.5 (Local Curvature-Dimension condition). We say that (X, d, m) satisfies $\text{CD}_{loc}(K, N)$ if and only if each point $x \in X$ has a neighborhood $X(x)$ such that for each pair $\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)$ supported in $X(x)$ there exists an optimal coupling π of $\mu_0 = \varrho_0 m$ and $\mu_1 = \varrho_1 m$, and a geodesic $\Gamma : [0, 1] \rightarrow \mathcal{P}_2(X, d, m)$ connecting μ_0 and μ_1 such that (2.3) holds true for all $t \in [0, 1]$ and all $N' \geq N$.

It is worth noticing that in the previous definition the geodesic Γ can exit from the neighborhood $X(x)$.

One of the main property of the reduced curvature dimension condition is the globalization one: under the non-branching assumption conditions $\text{CD}_{loc}^*(K, N)$ and $\text{CD}^*(K, N)$ are equivalent. Moreover it holds:

- $\text{CD}_{loc}^*(K, N)$ is equivalent to $\text{CD}_{loc}(K, N)$;
- $\text{CD}(K, N)$ implies $\text{CD}^*(K, N)$;
- $\text{CD}^*(K, N)$ implies $\text{CD}(K^*, N)$ where $K^* = K(N - 1)/N$.

Hence it is possible to pass from CD_{loc} to CD , but passing through CD^* and worsening the lower bound on the curvature. For all of these properties, see [4].

If a non-branching (X, d, m) satisfies $\text{CD}(K, N)$ then geodesics are unique $m \otimes m$ -a.e..

Lemma 2.6. Assume that (X, d, m) is non-branching and satisfies $\text{CD}(K, N)$ for some pair (K, N) . Then for every $x \in \text{supp}[m]$ and m -a.e. $y \in X$ (with the exceptional set depending on x) there exists a unique geodesic between x and y .

Moreover there exists a measurable map $\gamma : X^2 \rightarrow \mathcal{G}(X)$ such that for $m \otimes m$ -a.e. $(x, y) \in X^2$ the curve $t \mapsto \gamma_t(x, y)$ is the unique geodesic connecting x and y .

Under non-branching assumption is therefore possible to formulate $\text{CD}(K, N)$ in an equivalent point-wise version: (X, d, m) satisfies $\text{CD}(K, N)$ if and only if for each pair $\nu_0, \nu_1 \in \mathcal{P}_2(X, d, m)$ and each optimal coupling π of them

$$(2.4) \quad \varrho_t(\gamma_t(x_0, x_1)) \leq \left[\tau_{K, N'}^{(1-t)}(d(x_0, x_1)) \varrho_0^{-1/N'}(x_0) + \tau_{K, N'}^{(t)}(d(x_0, x_1)) \varrho_1^{-1/N'}(x_1) \right]^{-N},$$

for all $t \in [0, 1]$, and π -a.e. $(x_0, x_1) \in X \times X$. Here ϱ_t is the density of the push-forward of π under the map $(x_0, x_1) \mapsto \gamma_t(x_0, x_1)$.

We conclude with a partial list of properties enjoyed by metric measure spaces satisfying $\text{CD}^*(K, N)$ (or $\text{CD}_{loc}(K, N)$). If (X, d, m) verifies $\text{CD}^*(K, N)$ then:

- m is a doubling measure;
- m verifies Bishop-Gromov volume growth inequality;
- m verifies Brunn-Minkowski inequality;

with all of these properties stated in a quantitative form.

2.2. Spherical Hausdorff measure of codimension 1. What follows is contained in [3] and is valid under milder assumption than $\text{CD}^*(K, N)$ but for an easier exposition we assume (X, d, m) to satisfy $\text{CD}^*(K, N)$.

Let $B(X)$ be the set of balls and define the function $h : B(X) \rightarrow [0, \infty]$ as

$$h(\bar{B}_r(x)) := \frac{m(\bar{B}_r(x))}{r}.$$

Due to the doubling property of m , the function h turns out to be a doubling function, i.e. $h(\bar{B}_{2r}(x)) \leq (C_D/2)h(\bar{B}_r(x))$ for every $x \in X$, $r > 0$, where C_D is the doubling constant of m . Then, using the

Carathéodory construction, we may define the generalized Hausdorff spherical measure \mathcal{S}^h as

$$(2.5) \quad \mathcal{S}^h(A) := \liminf_{r \downarrow 0} \left\{ \sum_{i \in \mathbb{N}} h(B_i) : B_i \in \mathcal{B}(X), A \subset \bigcup_{i \in \mathbb{N}} B_i, \text{diam}(B_i) \leq r \right\}.$$

We will use the following property: given a set of finite perimeter E , the perimeter measure $P(E, \cdot)$ is absolutely continuous w.r.t. \mathcal{S}^h (see Theorem 4.4 of [3]).

2.3. Gradients and differentials. This part is taken from [11]. A curve $\gamma \in C([0, 1], X)$ is said to be absolutely continuous provided there exists $f \in L^1([0, 1])$ such that

$$d(\gamma_s, \gamma_t) \leq \int_s^t f(\tau) d\tau, \quad \forall s, t \in [0, 1], s \leq t.$$

Let $AC([0, 1], X)$ denote the set of absolutely continuous curves. If $\gamma \in AC([0, 1], X)$ then the limit

$$\lim_{\tau \rightarrow 0} \frac{d(\gamma_{t+\tau}, \gamma_t)}{\tau}$$

exists for a.e. $t \in [0, 1]$, is called metric derivative and denoted by $|\dot{\gamma}_t|$.

Given Borel functions $f : X \rightarrow \mathbb{R}, G : X \rightarrow [0, \infty]$ we say that G is an upper gradient of f provided

$$|f(\gamma_0) - f(\gamma_1)| \leq \int_0^1 G(\gamma_t) |\dot{\gamma}_t| dt, \quad \forall \gamma \in AC([0, 1], M),$$

where $|\dot{\gamma}_t|$ is the metric derivative of γ in t . For $f : X \rightarrow \mathbb{R}$ the local Lipschitz constant $|Df| : X \rightarrow [0, \infty]$ is defined by

$$|Df|(x) := \limsup_{y \rightarrow x} \frac{|f(y) - f(x)|}{d(y, x)}$$

if x is not isolated, and 0 otherwise. Define

$$|D^+ f|(x) := \limsup_{y \rightarrow x} \frac{(f(y) - f(x))^+}{d(y, x)}, \quad |D^- f|(x) := \limsup_{y \rightarrow x} \frac{(f(y) - f(x))^-}{d(y, x)},$$

the ascending and descending slope respectively. If f is locally Lipschitz, then $|D^\pm f|, |Df|$ are all upper gradients of f . In order to give a weaker notion of slope, consider the following family: $\gamma \in \mathcal{P}(C([0, 1], X))$ is a *test plan* if

$$e_{t\#} \gamma \leq Cm, \quad \forall t \in [0, 1], \quad \text{and} \quad \int_0^1 \int_0^1 |\dot{\gamma}_t| dt \gamma(d\gamma) < \infty,$$

where C is a positive constant. Therefore we have the following.

Definition 2.7. A Borel map $f : X \rightarrow \mathbb{R}$ belongs to the Sobolev class $S^2(X, d, m)$ (resp. $S_{loc}^2(X, d, m)$) if there exists a function $G \in L^2(X, m)$ (resp. $L_{loc}^2(X, m)$) such that

$$(2.6) \quad \int |f(\gamma_0) - f(\gamma_1)| \gamma(d\gamma) \leq \int \int_0^1 G(\gamma_s) |\dot{\gamma}_s| ds \gamma(d\gamma), \quad \forall \gamma \text{ test plan.}$$

If this is the case, G is called *weak upper gradient*.

For $f \in S^2(X, d, m)$ there exists a minimal function G , in the m -a.e. sense, in $L^2(X, m)$ such that (2.6) holds. Denote such minimal function with $|Df|_w$. Accordingly define the seminorm $\|f\|_{S^2(X, d, m)} := \| |Df|_w \|_{L^2(X, m)}$.

We now state a result on the weak upper gradient of Kantorovich potentials. Recall that $\gamma \in \mathcal{P}(\mathcal{G}(X))$ is a dynamical optimal plan if $\pi = (e_0, e_1)_\# \gamma \in \Pi(\mu_0, \mu_1)$ is optimal and the map $t \mapsto \mu_t := e_{t\#} \gamma$ is a geodesic in the 2-Wasserstein space.

Proposition 2.8 ([2], Theorem 10.3). *Let (X, d, m) verify $\text{CD}(K, N)$ for $K \in \mathbb{R}$ and $N \geq 1$. Let $\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)$, φ be a Kantorovich potential. Then for every γ optimal dynamical transference plan it holds*

$$d(\gamma_0, \gamma_1) = |D\varphi|_w(\gamma_0) = |D^+ \varphi|(\gamma_0), \quad \text{for } \gamma - \text{a.e. } \gamma.$$

If moreover the densities of μ_0 and of μ_1 are both in $L^\infty(M, m)$, then

$$\lim_{t \downarrow 0} \frac{\varphi(\gamma_t) - \varphi(\gamma_0)}{d(\gamma_0, \gamma_t)} = d(\gamma_0, \gamma_1), \quad \text{in } L^2(\mathcal{G}(M), \gamma).$$

In order to compute higher order derivatives, we introduce the following.

Definition 2.9. Let $f, g \in S^2(X, d, m)$. The functions

$$D^+f(\nabla g) := \liminf_{\varepsilon \downarrow 0} \frac{|D(g + \varepsilon f)|_w^2 - |Dg|_w^2}{2\varepsilon},$$

$$D^-f(\nabla g) := \limsup_{\varepsilon \uparrow 0} \frac{|D(g + \varepsilon f)|_w^2 - |Dg|_w^2}{2\varepsilon},$$

on $\{x \in M : |Dg|_w(x) \neq 0\}$, and are taken 0 by definition on $\{x \in M : |Dg|_w(x) = 0\}$.

Spaces where the two differentials coincide are called *infinitesimally strictly convex*, i.e. (X, d, m) is said to be infinitesimally strictly convex provided

$$(2.7) \quad \int D^+f(\nabla g)(x)m(dx) = \int D^-f(\nabla g)(x)m(dx), \quad \forall f, g \in S^2(X, d, m).$$

It is proven in [11] that (2.7) is equivalent to the point-wise one:

$$D^+f(\nabla g) = D^-f(\nabla g), \quad m - a.e., \quad \forall f, g \in S_{loc}^2(X, d, m).$$

If the space is infinitesimally strictly convex, we can denote by $Df(\nabla g)$ the common value. If the space is infinitesimally strictly convex, then $Df(\nabla g)$ is linear in f and 1-homogeneous and continuous in g .

There is a strong link between differentials and derivation along families of curves. For $\gamma \in \mathcal{P}(C([0, 1], X))$, define the norm $\|\gamma\|_2 \in [0, \infty]$ of γ by

$$\|\gamma\|_2^2 := \limsup_{t \downarrow 0} \frac{1}{t} \int_0^t |\dot{\gamma}_s|^2 ds \gamma(d\gamma),$$

if $\gamma \in \mathcal{P}(AC([0, 1], X))$ and $+\infty$ otherwise.

Definition 2.10. Let $g \in S^2(X, d, m)$. We say that $\gamma \in \mathcal{P}(C([0, 1], X))$ represents ∇g if γ is of bounded compression, $\|\gamma\|_2 < \infty$, and it holds

$$(2.8) \quad \liminf_{t \downarrow 0} \int \frac{g(\gamma_t) - g(\gamma_0)}{t} \gamma(d\gamma) \geq \frac{1}{2} (\|Dg|_w\|_{L^2(X, e_0 \# \gamma)}^2 + \|\gamma\|_2^2).$$

A straightforward consequence of (2.8) is that if γ represents ∇g , then the whole limit in the lefthand-side of (2.8) exists and verifies

$$\lim_{t \downarrow 0} \int \frac{g(\gamma_t) - g(\gamma_0)}{t} \gamma(d\gamma) = \frac{1}{2} (\|Dg|_w\|_{L^2(X, e_0 \# \gamma)}^2 + \|\gamma\|_2^2).$$

Theorem 2.11 ([11], Theorem 3.10). *Let $f, g \in S^2(X, d, m)$. For every $\gamma \in \mathcal{P}(C([0, 1], M))$ representing ∇g it holds*

$$\begin{aligned} \int D^+f(\nabla g)e_0 \# \gamma &\geq \limsup_{t \downarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \gamma(d\gamma) \\ &\geq \liminf_{t \downarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \gamma(d\gamma) \geq \int D^-f(\nabla g)e_0 \# \gamma. \end{aligned}$$

In order to have differentials enjoying more properties, we consider the following class of spaces:

Definition 2.12. We say that (X, d, m) is infinitesimally Hilbertian provided the seminorm $\|\cdot\|_{S^2(X, d, m)}$ on $S^2(X, d, m)$ satisfies the parallelogram rule.

We will use the following equivalent characterization: (X, d, m) is infinitesimally Hilbertian if and only if it is infinitesimally strictly convex and for every $f, g \in S_{loc}^2(X, d, m)$ it holds

$$(2.9) \quad Df(\nabla g) = Dg(\nabla f), \quad m - a.e..$$

Therefore in the case of infinitesimally Hilbertian spaces, $Df(\nabla g)$ will be linear in both f and g .

2.4. Hopf-Lax formula for Kantorovich potentials. What follows is contained in [2].

The definitions below make sense for a general Borel and real valued cost but we will only consider the d^2 case.

Definition 2.13. Let $\varphi : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$. Its d^2 -transform $\varphi^d : X \rightarrow \mathbb{R} \cup \{-\infty\}$ is defined by

$$\varphi^d(y) := \inf_{x \in X} \frac{1}{2} d^2(x, y) - \varphi(x).$$

Accordingly $\varphi : X \mapsto \mathbb{R} \cup \{\pm\infty\}$ is d^2 -concave if there exists $v : X \rightarrow \mathbb{R} \cup \{-\infty\}$ such that $\varphi = v^d$.

A d^2 -concave function φ such that (φ, φ^d) is a maximizing pair for the dual Kantorovich problem between μ_0, μ_1 is called a d^2 -concave Kantorovich potential for the couple (μ_0, μ_1) . A function φ is called a d^2 -convex Kantorovich potential if $-\varphi$ is a d^2 -concave Kantorovich potential.

We are interested in the evolution of potentials. They evolve accordingly to the Hopf-Lax evolution semigroup H_t^s via the following formula:

$$(2.10) \quad H_t^s(\psi)(x) := \begin{cases} \inf_{y \in X} \frac{1}{2} \frac{d^2(x, y)}{s-t} + \psi(y), & \text{if } t < s, \\ \psi(x), & \text{if } t = s, \\ \sup_{y \in X} \psi(y) - \frac{1}{2} \frac{d^2(x, y)}{t-s}, & \text{if } t > s. \end{cases}$$

We also introduce the rescaled cost $c^{t,s}$ defined by

$$c^{t,s}(x, y) := \frac{1}{2} \frac{d^2(x, y)}{s-t}, \quad \forall t < s, x, y \in X.$$

Observe that for $t < r < s$

$$c^{t,r}(x, y) + c^{r,s}(y, z) \geq c^{t,s}(x, z), \quad \forall x, y, z \in X,$$

and equality holds if and only if there is a constant speed geodesic $\gamma : [t, s] \rightarrow X$ such that $x = \gamma_t$, $y = \gamma_r$ and $z = \gamma_s$. The following result is taken from [15] (Theorem 7.36) but here we report a different version.

Theorem 2.14 ([1], Theorem 2.18). *Let $(\mu_t) \subset \mathcal{P}_2(X)$ be a constant speed geodesic in $(\mathcal{P}_2(X, d), d_W)$ and ψ a $c^{0,1}$ -convex Kantorovich potential for the couple (μ_0, μ_1) . Then $\psi_s := H_0^s(\psi)$ is a $c^{t,s}$ -concave Kantorovich potential for (μ_s, μ_t) , for any $t < s$.*

Similarly, if ϕ is a c -concave Kantorovich potential for (μ_1, μ_0) , then H_1^t is a $c^{t,s}$ -convex Kantorovich potential for (μ_t, μ_s) , for any $t < s$.

The following is an easy consequence.

Corollary 2.15. *Let φ be a d^2 -concave Kantorovich potential for μ_0, μ_1 and φ_t the Kantorovich potential for (μ_t, μ_1) where $\mu_t = e_{t\#}\gamma$ and γ is an optimal dynamical plan. Then the following holds:*

$$\varphi_t(\gamma_t) = \varphi(\gamma_0) - \frac{t}{2} d^2(\gamma_0, \gamma_1), \quad \gamma - \text{a.e. } \gamma.$$

Proof. Follows from Theorem 2.14 that

$$\varphi_t(x) = -H_1^t(\varphi^d)(x) = \inf_{y \in X} \frac{1}{2} \frac{d^2(x, y)}{1-t} - \varphi^d(y).$$

Therefore for γ -a.e. γ

$$\varphi_t(\gamma_t) \leq \frac{1}{2} \frac{d^2(\gamma_t, \gamma_1)}{1-t} + \varphi(\gamma_0) - \frac{1}{2} d^2(\gamma_0, \gamma_1) = \varphi(\gamma_0) - \frac{t}{2} d^2(\gamma_0, \gamma_1).$$

To prove the opposite inequality: observe that

$$\frac{d^2(\gamma_0, \gamma_t)}{t} + \frac{d^2(\gamma_t, y)}{1-t} \geq d^2(\gamma_0, y),$$

therefore for γ -a.e. γ

$$\frac{1}{2} \frac{d^2(\gamma_t, y)}{1-t} - \varphi^d(y) \geq \frac{1}{2} \frac{d^2(\gamma_t, y)}{1-t} - \frac{1}{2} d^2(\gamma_0, y) + \varphi(\gamma_0) \geq \varphi(\gamma_0) - \frac{1}{2} \frac{d^2(\gamma_0, \gamma_t)}{t} = \varphi(\gamma_0) - \frac{t}{2} d^2(\gamma_0, \gamma_1).$$

Taking the infimum the claim follows. \square

Also related to slopes of Kantorovich potentials is the following construction taken from [2]. Let $f : X \rightarrow \mathbb{R}$ be a Lipschitz function and consider the following maps

$$D^+(x, t) := \max \left\{ d(x, y) : y \in \operatorname{argmin} \left\{ y \mapsto f(y) + \frac{d^2(x, y)}{2t} \right\} \right\},$$

$$D^-(x, t) := \min \left\{ d(x, y) : y \in \operatorname{argmin} \left\{ y \mapsto f(y) + \frac{d^2(x, y)}{2t} \right\} \right\}.$$

Then a monotonicity result holds true.

Lemma 2.16 ([2], Proposition 3.1). *For all $x \in X$ it holds,*

$$D^+(x, t) \leq D^-(x, s), \quad 0 \leq t < s.$$

As a consequence, $D^+(x, \cdot)$ and $D^-(x, \cdot)$ are both nondecreasing, and they coincide with at most countably many exceptions in $[0, \infty)$.

2.5. Disintegration of measures. We conclude this introductory part with a short review on disintegration theory. What follows is taken from [6].

Given a measurable space (R, \mathcal{R}) and a function $r : R \rightarrow S$, with S generic set, we can endow S with the *push forward σ -algebra* \mathcal{S} of \mathcal{R} :

$$Q \in \mathcal{S} \iff r^{-1}(Q) \in \mathcal{R},$$

which could also be defined as the biggest σ -algebra on S such that r is measurable. Moreover given a measure space (R, \mathcal{R}, ρ) , the *push forward measure* η is then defined as $\eta := (r_# \rho)$.

Consider a probability space (R, \mathcal{R}, ρ) and its push forward measure space (S, \mathcal{S}, η) induced by a map r . From the above definition the map r is clearly measurable and inverse measure preserving.

Definition 2.17. A *disintegration* of ρ consistent with r is a map $\rho : \mathcal{R} \times S \rightarrow [0, 1]$ such that

- (1) $\rho_s(\cdot)$ is a probability measure on (R, \mathcal{R}) for all $s \in S$,
- (2) $\rho_s(B)$ is η -measurable for all $B \in \mathcal{R}$,

and satisfies for all $B \in \mathcal{R}, C \in \mathcal{S}$ the consistency condition

$$\rho(B \cap r^{-1}(C)) = \int_C \rho_s(B) \eta(ds).$$

A disintegration is *strongly consistent with respect to r* if for all s we have $\rho_s(r^{-1}(s)) = 1$.

The measures ρ_s are called *conditional probabilities*.

We say that a σ -algebra \mathcal{H} is *essentially countably generated* with respect to a measure m if there exists a countably generated σ -algebra $\hat{\mathcal{H}}$ such that for all $A \in \mathcal{H}$ there exists $\hat{A} \in \hat{\mathcal{H}}$ such that $m(A \triangle \hat{A}) = 0$.

We recall the following version of the disintegration theorem that can be found on [9], Section 452 (see [6] for a direct proof).

Theorem 2.18 (Disintegration of measures). *Assume that (R, \mathcal{R}, ρ) is a countably generated probability space, $R = \cup_{s \in S} R_s$ a partition of R , $r : R \rightarrow S$ the quotient map and (S, \mathcal{S}, η) the quotient measure space. Then \mathcal{S} is essentially countably generated w.r.t. η and there exists a unique disintegration $s \mapsto \rho_s$ in the following sense: if ρ_1, ρ_2 are two consistent disintegration then $\rho_{1,s}(\cdot) = \rho_{2,s}(\cdot)$ for η -a.e. s .*

If $\{S_n\}_{n \in \mathbb{N}}$ is a family essentially generating \mathcal{S} define the equivalence relation:

$$s \sim s' \iff \{s \in S_n \iff s' \in S_n, \forall n \in \mathbb{N}\}.$$

Denoting with p the quotient map associated to the above equivalence relation and with $(L, \mathcal{L}, \lambda)$ the quotient measure space, the following properties hold:

- $R_l := \cup_{s \in p^{-1}(l)} R_s = (p \circ r)^{-1}(l)$ is ρ -measurable and $R = \cup_{l \in L} R_l$;
- the disintegration $\rho = \int_L \rho_l \lambda(dl)$ satisfies $\rho_l(R_l) = 1$, for λ -a.e. l . In particular there exists a strongly consistent disintegration w.r.t. $p \circ r$;
- the disintegration $\rho = \int_S \rho_s \eta(ds)$ satisfies $\rho_s = \rho_{p(s)}$ for η -a.e. s .

In particular we will use the following corollary.

Corollary 2.19. *If $(S, \mathcal{S}) = (X, \mathcal{B}(X))$ with X Polish space, then the disintegration is strongly consistent.*

3. LEVEL SETS OF KANTOROVICH POTENTIALS

We fix here the objects that will be used throughout this note.

Let (X, d, m) be a non-branching metric measure space verifying $\text{CD}_{loc}(K, N)$ or equivalently $\text{CD}^*(K, N)$. Let $\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)$, $\pi \in \Pi(\mu_0, \mu_1)$ the optimal coupling and $\gamma \in \mathcal{P}(\mathcal{G}(X))$ an optimal dynamical transference plan such that $(e_0, e_1)_\# \gamma = \pi$. Moreover let φ be the corresponding d -concave Kantorovich potential for (μ_0, μ_1) and φ_t the d -concave Kantorovich potential for (μ_t, μ_1) . When it will be needed, we will use φ_0 instead of φ . From Corollary 2.15 it follows that $\varphi_1 = -\varphi^d$ μ_1 -a.e. and

$$\varphi_t(\gamma_t) = (1-t)\varphi_0(\gamma_0) + t\varphi_1(\gamma_1).$$

Fix the sets:

$$\Gamma := \left\{ (x, y) \in X \times X : \varphi(x) + \varphi^d(y) = \frac{d^2(x, y)}{2} \right\}, \quad \Gamma_a := \Gamma \cap \varphi^{-1}(a) \times X,$$

with $a \in \mathbb{R}$. It is well known that $\pi(\Gamma) = 1$. Fix also the support of γ , $G \subset \mathcal{G}(X)$ and the set of curves with starting point in $\varphi^{-1}(a)$:

$$G_a := \{\gamma \in G : \varphi(\gamma_0) = a\}.$$

The results of this section hold true even if μ_1 is singular with respect to m .

On the set $e_0(G)$ we will consider the partition given by the saturated sets of φ , i.e. $\{\varphi^{-1}(a)\}_{a \in \mathbb{R}}$. Disintegration Theorem implies that

$$m_{\mathbb{L}_{e_0}(G)} = \int_{\varphi(\mu_0)} \tilde{m}_a q(da), \quad \varphi_\#(m_{\mathbb{L}_{e_0}(G)}) = q,$$

with $\tilde{m}_a(\varphi^{-1}(a)^c) = 0$ for q -a.e. $a \in \varphi(\mu_0)$, where, in order to have a shorter notation, we have denoted by $\varphi(\mu_0)$ the set $\varphi(\text{supp}[\mu_0])$.

Proposition 3.1. *The measure $q = \varphi_\#(m_{\mathbb{L}_{e_0}(G)})$ is absolute continuous w.r.t. \mathcal{L}^1 .*

Proof. Step 1. Since we have to prove a local property, we can assume G to be compact and such that Proposition 2.8 holds true on it: $|D\varphi|_w(\gamma_0) = d(\gamma_0, \gamma_1)$, for all $\gamma \in G$. Define the map

$$e_0(G) \ni x \mapsto \hat{\varphi}(x) := \inf_{y \in e_1(G)} \left\{ \frac{d^2(x, y)}{2} - \varphi^d(y) \right\}.$$

Since G is compact, $e_0(G)$ and $e_1(G)$ are bounded and $\hat{\varphi}$ is obtained as the infimum of Lipschitz maps with uniformly bounded Lipschitz constant. Therefore $\hat{\varphi}$ is Lipschitz and coincide with $\varphi(x)$. Extend $\hat{\varphi}$ to the whole space keeping the same Lipschitz constant.

We can use the coarea formula (see Theorem 4.3 of [3]) in the particular case of Lipschitz maps: for any $B \subset X$ Borel

$$(3.1) \quad \int_{-\infty}^{+\infty} P(\{\hat{\varphi} = a\}, B) da \geq c \int_B \|\nabla \hat{\varphi}\|(x) m(dx),$$

where $c > 0$ depends on $\hat{\varphi}$, the modulus of the gradient of $\hat{\varphi}$ is defined as

$$\|\nabla \hat{\varphi}\|(x) := \liminf_{\rho \rightarrow 0} \frac{1}{\rho} \sup_{z \in B_\rho(x)} |\hat{\varphi}(z) - \hat{\varphi}(x)|$$

and $P(\{\hat{\varphi} = a\}, \cdot)$ is the perimeter measure.

Step 2. For $(x, y) \in (e_0, e_1)(G)$, $\|\nabla \hat{\varphi}\|(x) \geq d(x, y)$. Indeed fix $(x, y) \in (e_0, e_1)(G)$, then $\hat{\varphi}(x) + \varphi^d(y) = d^2(x, y)/2$ and

$$\hat{\varphi}(x) - \hat{\varphi}(z) \geq \frac{1}{2}(d^2(x, y) - d^2(z, y)) = \frac{1}{2}(d(x, y) - d(z, y))(d(x, y) + d(z, y))$$

Select a minimizing sequence $\rho_n \searrow 0$ for $\|\nabla \hat{\varphi}\|(x)$ and z_n on the geodesic connecting x to y at distance ρ_n from x . Then

$$\frac{1}{\rho_n} \sup_{z \in B_{\rho_n}(x)} |\hat{\varphi}(z) - \hat{\varphi}(x)| \geq \frac{1}{2} \frac{1}{\rho_n} (d(x, y) - d(z_n, y))(d(x, y) + d(z_n, y)) = \frac{1}{2}(d(x, y) + d(z_n, y)).$$

Passing to the limit we have $\|\nabla \hat{\varphi}\|(x) \geq d(x, y)$.

Let $E \subset \mathbb{R}$ with $\mathcal{L}^1(E) = 0$, then from (3.1) it follows that

$$\int_{\varphi^{-1}(E) \cap e_0(G)} \|\nabla \hat{\varphi}\|(x) m(dx) = \int_{\hat{\varphi}^{-1}(E) \cap e_0(G)} \|\nabla \hat{\varphi}\|(x) m(dx) \leq \frac{1}{c} \int_E P(\{\hat{\varphi} = a\}, e_0(G)) da = 0.$$

But on $e_0(G)$ the gradient of $\hat{\varphi}$ is strictly positive, it follows that $m(\varphi^{-1}(E) \cap e_0(G)) = 0$ and therefore the claim is proved. \square

Proposition 3.1 implies that

$$m_{\perp e_0(G)} = \int_{\varphi(\mu_0)} \tilde{m}_a q(a) \mathcal{L}^1(da) = \int_{\varphi(\mu_0)} \hat{m}_a \mathcal{L}^1(da).$$

An analogous partition can be considered also on the support of μ_t , $e_t(G)$. Indeed the d^2 -cyclical monotonicity of Γ implies that the family

$$\{\gamma_t : \varphi(\gamma_0) = a\}_{a \in \varphi(\mu_0)}$$

is indeed a disjoint family and a partition of $e_t(G)$. We can therefore consider for every $t \in [0, 1]$ the disintegration of $m_{\perp e_t(G)}$ w.r.t. the aforementioned family. Since for every $t \in [0, 1]$

$$\mu_0(\varphi^{-1}(A)) = \mu_t(\{\gamma_t : \varphi(\gamma_0) \in A\}),$$

the quotient measures of μ_0 and μ_t are the same measure. Hence the quotient measure of $m_{\perp e_t(G)}$ and the one of $m_{\perp e_0(G)}$ are equivalent. It follows that

$$(3.2) \quad m_{\perp e_t(G)} = \int_{\varphi(\mu_0)} \tilde{m}_{a,t} f_t(a) \mathcal{L}^1(da) = \int_{\varphi(\mu_0)} \hat{m}_{a,t} \mathcal{L}^1(da), \quad \hat{m}_{a,t}(\{\gamma_t : \varphi(\gamma_0) = a\}^c) = 0.$$

To keep notation consistent, we will denote also the conditional probabilities for $t = 0$ with $\hat{m}_{a,0}$.

For $t \in (0, 1)$, still the partition is formed by saturated sets of a suitable map Φ_t :

$$(3.3) \quad e_t(G) \ni \gamma_t \mapsto \Phi_t(\gamma_t) := \varphi_t(\gamma_t) + \frac{t}{2} d^2(\gamma_0, \gamma_1),$$

where in the definition of Φ_t we used that, for $t \in (0, 1)$, for every $x \in e_t(G)$ there exists only one geodesic $\gamma \in G$ with $\gamma_t = x$. It is straightforward from Corollary 2.15 to observe that $\{\gamma_t : \varphi(\gamma_0) = a\} = \Phi_t^{-1}(a)$. It follows that for $t \neq 1$ the conditional probabilities are all of the same type.

Lemma 3.2. *For every $t \in [0, 1]$ and \mathcal{L}^1 -a.e. $a \in \varphi(\mu_0)$, with the exceptional set depending on t , it holds*

$$\hat{m}_{a,t} \ll \mathcal{S}^h,$$

where \mathcal{S}^h is the spherical Hausdorff measure of codimension one defined in (2.5).

Proof. Step 1. For $t = 0$ the claim follows from the proof of Proposition 3.1, where we have shown that $\hat{m}_a \leq P(\{\hat{\varphi} = a\}, \cdot)$ the latter being absolutely continuous with respect to \mathcal{S}^h ([3], Theorem 4.3).

Fix $t \in (0, 1)$. The map Φ_t is Lipschitz, indeed as already observed in the proof of Proposition 3.1, the Kantorovich potential φ_t can be assumed to be Lipschitz and

$$d^2(\gamma_t, \gamma_1) = 2(1-t) \sup_{y \in X} \{\varphi_t(\gamma_t) + \varphi^d(y)\}.$$

Therefore Φ_t is Lipschitz on $e_t(G)$. Denote with $\hat{\Phi}_t$ its Lipschitz extension to the whole X . Coarea formula for Lipschitz maps applies (see Theorem 4.3 of [3]): for any measurable $A \subset X$

$$\int_{-\infty}^{+\infty} P(\{\hat{\Phi}_t = a\}, A) \mathcal{L}^1(da) \geq c \int_A \|\nabla \hat{\Phi}_t\|(x) m(dx),$$

where c is a strictly positive constant.

Step 2. Since $\hat{\Phi}_t$ is Lipschitz, Theorem 4.4 and Theorem 4.6 of [3] imply that

$$P(\{\hat{\Phi}_t = a\}, \cdot) \leq C_N \mathcal{S}^h.$$

where C_N is the doubling constant of m . So we have

$$\int \|\nabla \hat{\Phi}_t\| \hat{m}_{a,t} \mathcal{L}^1(da) \leq \int_{-\infty}^{+\infty} P(\{\hat{\Phi}_t = a\}, \cdot) \mathcal{L}^1(da),$$

which in turn gives

$$\|\nabla \hat{\Phi}_t\|_{\hat{m}_{a,t}} \leq C_N \mathcal{S}^h.$$

Since the set $\{x \in e_t(G) : \|\nabla \Phi_t\|(x) = 0\} \subset \Phi_t^{-1}(N)$ for some $N \subset \mathbb{R}$ with $\mathcal{L}^1(N) = 0$, it follows that $m(\{x \in e_t(G) : \|\nabla \Phi_t\|(x) = 0\}) = 0$ and the claim follows. \square

Remark 3.3. The statement of Lemma 3.2 is true also for $t = 1$ only if $\mu_1 \ll m$. Indeed if that is the case, there exists a unique optimal map from μ_0 to μ_1 and this map is invertible, see [10], Theorem 2.7. Hence $d^2(\gamma_0, \gamma_1)$ depends only on γ_1 and since

$$d^2(\gamma_0, \gamma_1) = 2 \sup_{x \in X} \{\varphi_0(x) + \varphi^d(\gamma_1)\},$$

it is also Lipschitz. Then following the proof of Lemma 3.2, we get the statement also for the case $t = 1$.

4. d -CYCLICAL MONOTONICITY

In this section a is a fixed element of $\varphi(\mu_0)$. We show that a d^2 -cyclically monotone transport plan restricted to the level sets of the potential φ is d -cyclically monotone. We recall here the definition $\Gamma_a := \Gamma \cap \varphi^{-1}(a) \times X$.

Lemma 4.1. *The set Γ_a is d -cyclically monotone.*

Proof. Let $(x_i, y_i) \in \Gamma_a$ for $i = 1, \dots, n$ and observe that

$$\frac{1}{2}d^2(x_i, y_i) = \varphi(x_i) + \varphi^d(y_i) = \varphi(x_{i-1}) + \varphi^d(y_i) \leq \frac{1}{2}d^2(x_{i-1}, y_i).$$

Hence $d(x_{i-1}, y_i) \geq d(x_i, y_i)$ and therefore

$$\sum_{i=1}^n d(x_i, y_{i+1}) \geq \sum_{i=1}^n d(x_i, y_i)$$

and the claim follows. \square

For $r \in [0, 1]$ define the “closed” and “open” evolution sets as follows

$$\bar{\Gamma}_a(r) := \{\gamma_t : (\gamma, t) \in G_a \times [0, r]\}, \quad \Gamma_a(r) := \{\gamma_t : (\gamma, t) \in G_a \times [0, r)\},$$

where $G_a = G \cap (\varphi \circ e_0)^{-1}(a)$.

Lemma 4.2. *The family $\{e_t(G_a)\}_{t \in [0, 1]}$ is a partition of $\bar{\Gamma}_a(1)$.*

Proof. By construction the family covers $\bar{\Gamma}_a(1)$, so we have only to show that overlapping doesn't occur. Assume by contradiction the existence of $\hat{\gamma}, \tilde{\gamma} \in G_a$, $\hat{\gamma} \neq \tilde{\gamma}$ such that $\hat{\gamma}_s = \tilde{\gamma}_t = z$ with, say, $s < t$.

Then d -cyclical monotonicity and non-branching property of (X, d, m) imply that $\hat{\gamma}$ and $\tilde{\gamma}$ form a cycle of zero cost (they are contained in a longer geodesic): if $\hat{\gamma}_0 = x_0, \hat{\gamma}_1 = y_0$ and $\tilde{\gamma}_0 = x_1, \tilde{\gamma}_1 = y_1$ then

$$d(x_0, y_1) + d(x_1, y_0) = d(x_0, y_0) + d(x_1, y_1).$$

There are two possible cases: or $d(x_1, y_0) \leq d(x_0, y_0)$ or $d(x_0, y_1) \leq d(x_1, y_1)$, indeed if both were false we would have a contradiction with the previous identity. In the first case

$$\frac{1}{2}d^2(x_0, y_0) = \varphi(x_0) + \varphi^d(y_0) = \varphi(x_1) + \varphi^d(y_0) \leq \frac{1}{2}d^2(x_1, y_0) \leq \frac{1}{2}d^2(x_0, y_0).$$

Therefore $d(x_0, y_0) = d(x_1, y_0)$ and since they lie on the same geodesic $x_0 = x_1$. In the second case

$$\frac{1}{2}d^2(x_1, y_1) = \varphi(x_1) + \varphi^d(y_1) = \varphi(x_0) + \varphi^d(y_1) \leq \frac{1}{2}d^2(x_0, y_1) \leq \frac{1}{2}d^2(x_1, y_1),$$

and the same conclusion holds true: $x_0 = x_1$.

Hence we have $(x_0, y_0), (x_0, y_1) \in \Gamma_a$. It follows from Proposition 2.8 that for all $\gamma \in G$

$$|D\varphi|_w(x) = d(\gamma_0, \gamma_1).$$

Therefore necessarily $y_0 = y_1$. Since $\hat{\gamma}, \tilde{\gamma}$ have also an inner common point, they must coincide implying a contradiction. \square

Following Lemma 4.2, we can consider the disintegration of $m_{\perp \bar{\Gamma}_a(1)}$ w.r.t. the family of sets $\{e_t(G_a)\}_{t \in [0,1]}$:

$$m_{\perp \bar{\Gamma}_a(1)} = \int_{[0,1]} \bar{m}_{a,t} q_a(dt), \quad q_a \in \mathcal{P}([0,1]), \quad q_a(I) = m(\varphi^{-1}(I) \cap \bar{\Gamma}_a(1)).$$

Observe that any $\gamma \in G_a$ can be taken as quotient set, therefore Corollary 2.19 implies the strong consistency of the disintegration, i.e. for q_a -a.e. $t \in [0,1]$ $\bar{m}_{a,t}$ is concentrated on $e_t(G_a)$.

Proposition 4.3. *The quotient measure q_a is absolute continuous with respect to \mathcal{L}^1 .*

Proof. Since Γ_a is d -cyclically monotone, we can consider another partition of $\bar{\Gamma}_a(1)$. Consider the family of sets $\{\gamma_s : s \in [0,1]\}_{\gamma \in G_a}$. Even if this family is not made of disjoint sets, notice that possible intersection may occur only on $e_0(G_a)$ and $e_1(G_a)$. Since $m(e_0(G_a)) = m(e_1(G_a)) = 0$, we have that

- the following disintegration holds true:

$$m_{\perp \bar{\Gamma}_a(1)} = \int \eta_y q_\gamma(dy),$$

where the quotient measure q_γ is concentrated on $\{\gamma_{1/2} : \gamma \in G_a\}$ and q_γ -a.e. conditional probability η_y is concentrated on $\{\gamma_s : s \in [0,1], \gamma \in e_{1/2}^{-1}(y) \cap G_a\}$;

- Since $\text{CD}_{loc}(K, N)$ implies $\text{MCP}(K, N)$, from Theorem 9.5 of [7] we have that $\eta_y = g(y, \cdot) \mathcal{L}^1_{[0,1]}$ for q_γ -a.e. y , and for $r \leq R$

$$(4.1) \quad \left(\frac{\sin\left(\frac{r}{R}d(\gamma_0, \gamma_R)\sqrt{K/(N-1)}\right)}{\sin\left(d(\gamma_0, \gamma_R)\sqrt{K/(N-1)}\right)} \right)^{N-1} \leq \frac{g(y, r)}{g(y, R)} \leq \left(\frac{\sin\left(\frac{r}{R}d(\gamma_r, \gamma_1)\sqrt{K/(N-1)}\right)}{\sin\left(d(\gamma_r, \gamma_1)\sqrt{K/(N-1)}\right)} \right)^{N-1},$$

where $\gamma = e_{1/2}^{-1}(y) \cap G_a$.

To prove the claim it is enough to observe that the two disintegration proposed for $m_{\perp \bar{\Gamma}_a(1)}$ are the same. Use Fubini's Theorem to get

$$\int_{[0,1]} \bar{m}_{a,t} q_a(dt) = m_{\perp \bar{\Gamma}_a(1)} = \int g(y, \cdot) \mathcal{L}^1(dt) q_\gamma(dy) = \int_{[0,1]} g(\cdot, t) q_\gamma(dy) dt,$$

therefore from uniqueness of disintegration, $\bar{m}_{a,t} = g(\cdot, t) q_\gamma \left(\int g(y, t) q_\gamma(dy) \right)^{-1}$ and $q_a = \left(\int g(y, t) q_\gamma(dy) \right) \mathcal{L}^1$. \square

Hence if $dq_a/d\mathcal{L}^1$ denotes the density of q_a with respect to \mathcal{L}^1 , posing $m_{a,t} := (dq_a/d\mathcal{L}^1) \bar{m}_{a,t}$, we have

$$(4.2) \quad m_{\perp \bar{\Gamma}_a(1)} = \int_{[0,1]} m_{a,t} dt.$$

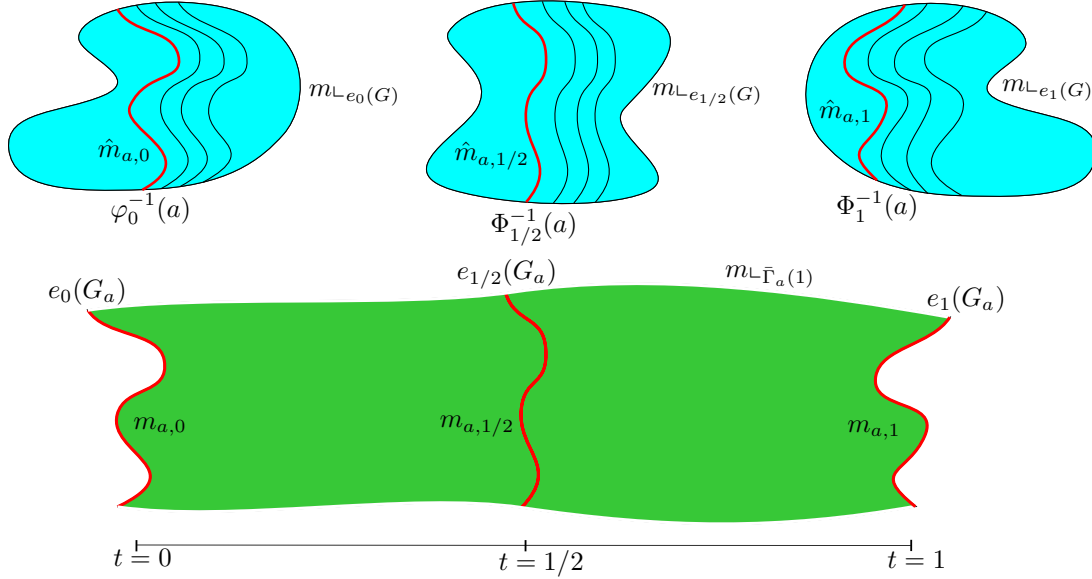
Denote by $\|m_{a,t}\| := m_{a,t}(e_t(G_a))$. During the proof of Proposition 4.3 we have also found an estimate on $\|m_{a,t}\|$. Indeed integrating (4.1) in y , we get

$$(4.3) \quad \inf_{\gamma \in G_a} \left(\frac{\sin\left(\frac{r}{R}d(\gamma_0, \gamma_R)\sqrt{\frac{K}{(N-1)}}\right)}{\sin\left(d(\gamma_0, \gamma_R)\sqrt{\frac{K}{(N-1)}}\right)} \right)^{N-1} \leq \frac{\|m_{a,r}\|}{\|m_{a,R}\|} \leq \sup_{\gamma \in G_a} \left(\frac{\sin\left(\frac{r}{R}d(\gamma_r, \gamma_1)\sqrt{\frac{K}{(N-1)}}\right)}{\sin\left(d(\gamma_r, \gamma_1)\sqrt{\frac{K}{(N-1)}}\right)} \right)^{N-1}.$$

Remark 4.4. Here we want to stress the differences between $\hat{m}_{a,t}$ and $m_{a,t}$. It is worth underlining again that both measures are concentrated on $e_t(G_a)$. Also they are both obtained as conditional measures of m or, otherwise stated, they belong to the range of two different disintegration maps of m :

$$m_{\perp e_t(G)} = \int \hat{m}_{a,t} \mathcal{L}^1(da), \quad m_{\perp \bar{\Gamma}_a(1)} = \int m_{a,t} \mathcal{L}^1(dt).$$

Since in both disintegrations the quotient measure is \mathcal{L}^1 , conditional measures can be interpreted as the “derivative” with respect to the parameter in the quotient space, a in the first case and t in the second one, of m . Even if m and $e_t(G_a)$ are fixed, what matters, and implies $\hat{m}_{a,t} \neq m_{a,t}$, is the difference between $e_{t+\varepsilon}(G_a)$ and $e_t(G_{a+\varepsilon})$. The difference can be observed in Figure 1.


 FIGURE 1. Above and below the disintegration with conditional $\hat{m}_{a,t}$ and $m_{a,t}$, respectively.

5. UNIQUENESS OF CONDITIONAL MEASURES

It is necessary to find a comparison between $m_{a,t}$ and $\hat{m}_{a,t}$ (note that they are both concentrated on $e_t(G_a)$). Indeed disintegrate γ w.r.t. the saturated sets of $\varphi \circ e_0$:

$$\gamma = \int_{\varphi(\mu_0)} \gamma_a q(a) \mathcal{L}^1(da), \quad \gamma_a((\varphi \circ e_0)^{-1}(a)) = 1.$$

The quotient measure $q(a) \mathcal{L}^1(da)$ is the same quotient measure of μ_t for every $t \in [0, 1]$. Then necessarily, from uniqueness of disintegration,

$$e_{t\#} \gamma_a = \left(\int \varrho_t(z) \hat{m}_{a,t}(dz) \right)^{-1} \varrho_t \hat{m}_{a,t}.$$

On the other hand, in order to apply $\text{CD}^*(K, N)$ or $\text{CD}_{loc}(K, N)$ to $e_{t\#} \gamma_a$, it seems to be almost mandatory to express $e_{t\#} \gamma_a$ in terms of $m_{a,t}$. Therefore the scope of this section is to show that $m_{a,t}$ and $\hat{m}_{a,t}$ are not singular with respect to each other and to write the mutual densities.

Lemma 5.1. *For every $a \in \varphi(\mu_0)$,*

$$\lim_{s \rightarrow 0} \frac{1}{s} \int_{(t, t+s)} m_{a,\tau} \mathcal{L}^1(d\tau) = m_{a,t},$$

for \mathcal{L}^1 -a.e. $t \in [0, 1]$, where the convergence is in the weak sense.

Proof. Since (X, d, m) is locally compact, the space of real valued continuous and bounded functions $C_b(X)$ is separable. Let $\{f_k\}_{k \in \mathbb{N}} \subset C_b(X)$ be a dense family.

Fix $a \in \varphi(\mu_0)$. The Lebesgue differentiation theorem implies that for every $k \in \mathbb{N}$

$$\frac{1}{s} \int_{(t, t+s)} \left(\int f_k(z) m_{a,\tau}(dz) \right) \mathcal{L}^1(d\tau) \rightarrow \int f_k(z) m_{a,t}(dz), \quad \text{as } s \searrow 0,$$

as real numbers, for all $t \in [0, 1] \setminus E_{a,k}$ with $\mathcal{L}^1(E_{a,k}) = 0$. Hence $E_a := \cup_{k \in \mathbb{N}} E_{a,k}$ is \mathcal{L}^1 -negligible. Take $f \in C_b(X)$ and chose $\{f_{k_h}\}_{h \in \mathbb{N}}$ approximating f in the uniform norm. Using f_{k_h} , it is then fairly easy to show that

$$\lim_{s \rightarrow 0} \frac{1}{s} \int_{(t, t+s)} \left(\int f(z) m_{a,\tau}(dz) \right) \mathcal{L}^1(d\tau) = \int f(z) m_{a,t}(dz)$$

for all $t \in [0, 1] \setminus E_a$. □

The analogous statement of Lemma 5.1 is true for the conditional measures $\hat{m}_{a,t}$ of (3.2): fix $t \in [0, 1]$, then

$$\lim_{b \rightarrow 0} \frac{1}{b} \int_{(a, a+b)} \hat{m}_{a,t} \mathcal{L}^1(d\alpha) = \hat{m}_{a,t},$$

for \mathcal{L}^1 -a.e. $a \in \varphi(\mu_0)$, where the convergence is in the weak sense.

The next one is the main technical statement of the section.

Proposition 5.2. *For \mathcal{L}^2 -a.e. $(a, t) \in \varphi(\mu_0) \times [0, 1]$ the two sequences*

$$\left\{ n \cdot m_{\Phi_t^{-1}([a-1/n, a])} \right\}_{n \in \mathbb{N}}, \quad \left\{ n \cdot m_{\Phi_t^{-1}([a-1/n, a]) \cap \bar{\Gamma}_a(1)} \right\}_{n \in \mathbb{N}}$$

have the same weak limit, where Φ_t was introduced in (3.3) and verifies $\Phi_t(\gamma_t) = \varphi_0(\gamma_0)$.

Proof. Step 1. Fix $a \in \varphi(\mu_0)$ and consider the map

$$[0, 1] \ni t \mapsto \mathcal{S}^h_{\lfloor e_t(G_a)} \in \mathcal{M}(X, d),$$

where $\mathcal{M}(X, d)$ is the set of finite Borel measures over X endowed with the weak topology, indeed $\mathcal{S}^h_{\lfloor e_t(G_a)}$ is finite for a.e. $t \in [0, 1]$. The map is \mathcal{L}^1 -measurable, therefore for every $\varepsilon > 0$ there exists a compact set $I \subset [0, 1]$ with $\mathcal{L}^1(I) \geq 1 - \varepsilon$ such that the map

$$I \ni t \mapsto \mathcal{S}^h_{\lfloor e_t(G_a)}$$

is continuous.

Step 2. Observe that the claim is equivalent to

$$\limsup_{n \rightarrow \infty} n \cdot m(\Phi_t^{-1}([a - 1/n, a]) \setminus \bar{\Gamma}_a(1)) = 0.$$

So suppose by contradiction the existence of a sequence n_k such that

$$n_k \int_{[a-1/n_k, a]} \hat{m}_{b,t}(\bar{\Gamma}_a(1)^c) \mathcal{L}^1(db) \geq \alpha, \quad \text{for } k \text{ sufficiently large.}$$

In Lemma 3.2 we have proved that

$$\|\nabla \hat{\Phi}_t\| \hat{m}_{a,t} \leq C_N \mathcal{S}^h,$$

with $\|\nabla \hat{\Phi}_t\|$ positive m -a.e.. Then necessarily there exists a sequence $\{a_k\}_{k \in \mathbb{N}}$ converging to a , with $a_k \in [a - 1/n_k, a]$, such that $\hat{m}_{a_k,t}(\bar{\Gamma}_a(1)^c) \geq \alpha'$ and $\mathcal{S}^h(\bar{\Gamma}_a(1)^c \cap \Phi_t^{-1}(a_k)) \geq \alpha'$.

For $k \in \mathbb{N}$ sufficiently large, let $H_k \subset \bar{\Gamma}_a(1)^c \cap \Phi_t^{-1}(a_k)$ be a compact set such that $\mathcal{S}^h(H_k) \geq \alpha'/2$. Existence of H_k follows from inner regularity. Being each $H_k \subset e_t(G_{a_k}) \subset e_t(G)$, the sequence $\{H_k\}_{k \in \mathbb{N}}$ is precompact in Hausdorff topology, hence possibly passing to subsequences H_k is converging to $H \subset e_t(G_a)$. Moreover, possibly choosing a smaller set, we also have $\mathcal{S}^h(H_k) \leq C$. Then from tightness, passing again to subsequences

$$\mathcal{S}^h_{\lfloor H_k} \rightharpoonup \eta \in \mathcal{M}(X, d),$$

with $\eta(H) \geq \limsup_k \mathcal{S}^h(H_k) \geq \alpha'/2$.

Step 3. For every $\varepsilon' > 0$ there exists a family $\{B_i\}_{i \in \mathbb{N}} \subset B(X)$ such that

$$\mathcal{S}^h(e_t(G_a)) \geq \sum_{i=1}^{\infty} h(B_i) - \varepsilon', = \sum_{i=1}^{\infty} \frac{m(B_i)}{r_i} - \varepsilon'.$$

where r_i is the radius of the ball B_i . We want to consider just finitely many B_i : since the summation is finite, for every $\delta > 0$ there exists $n_\delta \in \mathbb{N}$ such that

$$\sum_{i=n_\delta}^{\infty} h(B_i) \leq \delta \mathcal{S}^h(e_t(G_a)).$$

Denote by $A_\delta = \cup_{i \leq n_\delta} B_i$ and consider any sequence $\{t_k\}_{k \in \mathbb{N}} \subset I$ such that $t_k \rightarrow t$. Then we have

$$\begin{aligned} \mathcal{S}^h(e_t(G_a)) &\geq \sum_{i=1}^{n_\delta} h(B_i) - \varepsilon' \\ &\geq \mathcal{S}^h(A_\delta \cap (e_{t_k}(G_a) \cup H_k)) - \varepsilon' \\ &= \mathcal{S}^h(A_\delta \cap e_{t_k}(G_a)) + \mathcal{S}^h(A_\delta \cap H_k) - \varepsilon' \\ &= \mathcal{S}^h \llcorner_{e_{t_k}(G_a)}(A_\delta) + \mathcal{S}^h \llcorner_{H_k}(A_\delta) - \varepsilon' \\ &\geq \mathcal{S}^h \llcorner_{e_t(G_a)}(A_\delta) + \eta(A_\delta) - \varepsilon' \\ &\geq (1 - \delta) \mathcal{S}^h(e_t(G_a)) + \eta(A_\delta) - \varepsilon', \end{aligned}$$

where between the second and third line we have used $e_{t_k}(G_a) \subset \Gamma_a(1)$ while $H_k \cap \Gamma_a(1) = \emptyset$, and from the fifth to the sixth we used lower semicontinuity on open sets of weakly converging measures.

Hence for every $\varepsilon' > 0$ and $\delta > 0$

$$\mathcal{S}^h(e_t(G_a)) \geq (1 - \delta) \mathcal{S}^h(e_t(G_a)) + \eta(A_\delta) - \varepsilon'.$$

Letting $\delta \rightarrow 0$, n_δ goes to ∞ , hence

$$\mathcal{S}^h(e_t(G_a)) \geq \mathcal{S}^h(e_t(G_a)) + \eta(\cup_{i=1}^\infty B_i) - \varepsilon'.$$

Recall that $\{B_i\}_{i \in \mathbb{N}}$ is a covering of $e_t(G_a)$ and $H \subset e_t(G_a)$, therefore

$$\mathcal{S}^h(e_t(G_a)) \geq \mathcal{S}^h(e_t(G_a)) + \eta(H) - \varepsilon' \geq \mathcal{S}^h(e_t(G_a)) + \alpha'/2 - \varepsilon'.$$

Since ε' can be as small as we need, we have a contradiction and the claim is proved for $a \in \varphi(\mu_0)$ fixed and $t \in I$. Lusin Theorem implies the statement. \square

Since $n \cdot m_{\Phi_t^{-1}([a-1/n, a])}$ weakly converges to $\hat{m}_{a,t}$, Theorem ?? suggests to study the following quantity: for fixed $x \in e_t(G)$

$$(5.1) \quad \lambda_t(x) := \lim_{n \rightarrow \infty} n \cdot \mathcal{L}^1 \left(\left\{ \tau \in [0, 1] : \gamma_\tau \in \Phi_t^{-1}([\varphi(\gamma_0) - 1/n, \varphi(\gamma_0)]) \right\} \right),$$

where γ is the unique geodesic in G such that $\gamma_t = x$. Before proving that λ_t is the correct density, we need few technical results.

Lemma 5.3. *For γ -a.e. $\gamma \in G$, $\Phi_t(\gamma_t) \geq \Phi_t(\gamma_{t+s})$ for \mathcal{L}^1 -a.e. $s \in [0, \varepsilon]$, for a suitable $\varepsilon > 0$ depending on γ .*

Proof. Observe that

$$d^2((e_0, e_1)(e_t^{-1}(\gamma_{t+s}))) = \frac{1}{(1-t)^2} d^2(\gamma_{t+s}, \gamma_1^s),$$

where $\gamma^s = e_t^{-1}(\gamma_{t+s})$. Moreover

$$d^2(\gamma_{t+s}, \gamma_1^s) = D^2(\gamma_{t+s}, 1-t),$$

where $D(x, t) = d(x, y)$ for $y \in \operatorname{argmin}\{y \mapsto -\varphi^d(y) + \frac{d^2(x, y)}{2t}\}$. Note that non-branching implies that D is a well defined map. From Lemma 2.16, $D(x, \cdot)$ is nondecreasing, therefore

$$D^2(\gamma_{t+s}, 1-t) \geq D^2(\gamma_{t+s}, 1-(t+s)) = d^2(\gamma_{t+s}, \gamma_1) = \frac{(1-(t+s))^2}{(1-t)^2} d^2(\gamma_t, \gamma_1).$$

Hence

$$d^2((e_0, e_1)(e_t^{-1}(\gamma_{t+s}))) \geq (1-(t+s))^2 d^2(\gamma_0, \gamma_1).$$

Starting from 1 instead of 0, from Lemma 2.16 we can deduce reasoning as before that

$$d^2((e_0, e_1)(e_t^{-1}(\gamma_{t+s}))) \leq \left(1 + \frac{s}{t}\right)^2 d^2(\gamma_0, \gamma_1).$$

So we have that $s \mapsto d^2(\gamma_0^s, \gamma_1^s)$ is Lipschitz. Moreover it is well-known that

$$\lim_{s \rightarrow 0} \frac{\varphi_t(\gamma_{t+s}) - \varphi_t(\gamma_t)}{s} = -d^2(\gamma_0, \gamma_1).$$

It is then straightforward to check that

$$\limsup_{s \rightarrow 0} \frac{\Phi_t(\gamma_{t+s}) - \Phi_t(\gamma_t)}{s} \leq -d^2(\gamma_0, \gamma_1) + \frac{t}{2} \frac{2}{t} d^2(\gamma_0, \gamma_1) = 0,$$

and the claim follows. \square

We deduce the following.

Corollary 5.4. *The quantity $\lambda_t(x)$ is well defined and verifies*

$$\frac{1}{\lambda_t(\gamma_t)} = \liminf_{\tau \rightarrow 0} \frac{\Phi_t(\gamma_t) - \Phi_t(\gamma_{t+\tau})}{\tau}.$$

Proof. Call $a = \varphi(\gamma_0)$. There is only one $\tau \in [0, 1]$ such that $\Phi_t(\gamma_\tau) = a$ and it must be t . Indeed if the converse would be true, then we would have two points on the geodesic γ coming from two distinct geodesics both starting from $\varphi^{-1}(a)$, contradicting the d -cyclical monotonicity. It follows that for n large enough $(\Phi_t \circ \gamma)^{-1}([a - 1/n, a])$ is a closed interval converging to t as $n \nearrow \infty$. Hence for n suitably large

$$\mathcal{L}^1\left(\left\{\tau \in [0, 1] \cap (\Phi_t \circ \gamma_{(\cdot)})^{-1}([\varphi(\gamma_0) - 1/n, \varphi(\gamma_0)])\right\}\right) = \tau_n,$$

where $t + \tau_n = \max\{\tau : \Phi_t(\gamma_\tau) = a - 1/n\}$.

Take any sequence $\tilde{\tau}_n \rightarrow 0$. Since all the previous results hold true if we replace $a - 1/n$ with any sequence $a - \varepsilon_n$ with $\varepsilon_n \searrow 0$ as $n \nearrow \infty$, without loss of generality we can assume that $\Phi_t(\gamma_{t+\tilde{\tau}_n}) = a - 1/n$. It follows

$$\frac{\Phi_t(\gamma_t) - \Phi_t(\gamma_{t+\tilde{\tau}_n})}{\tilde{\tau}_n} \geq \frac{\Phi_t(\gamma_t) - \Phi_t(\gamma_{t+\tau_n})}{\tau_n} = \frac{1}{n \mathcal{L}^1\left(\left\{\tau \in [0, 1] \cap (\Phi_t \circ \gamma_{(\cdot)})^{-1}([a - 1/n, a])\right\}\right)}.$$

Hence the limit in (5.1), being equal to a liminf is well defined and identity of the claim is proved. \square

The following is the main result of the section.

Theorem 5.5. *For γ -a.e. $\gamma \in G$ and \mathcal{L}^1 -a.e. $t \in [0, 1]$*

$$\hat{m}_{a,t} = \lambda_t m_{a,t}.$$

Proof. Step 1. For \mathcal{L}^2 -a.e. $(a, t) \in \varphi(\mu_0) \times [0, 1]$ Theorem ?? applies:

$$\hat{m}_{a,t} = \lim_{n \rightarrow \infty} n \cdot m_{\mathbb{L}_{\Phi_t^{-1}([a-1/n, a]) \cap \bar{\Gamma}_a(1)}}.$$

Moreover, as we saw in the proof of Proposition 4.3, $m_{\mathbb{L}_{\bar{\Gamma}_a(1)}} = \int g(y, \cdot) \mathcal{L}^1 q(dy)$, hence

$$m_{\mathbb{L}_{\Phi_t^{-1}([a-1/n, a]) \cap \bar{\Gamma}_a(1)}} = \int_S \left(g(y, \cdot) \mathcal{L}^1_{\mathbb{L}_{(\Phi_t \circ \gamma(y))^{-1}([a-1/n, a])}} \right) q(dy).$$

Multiply and divide by $\mathcal{L}^1\left(\left\{\tau \in [0, 1] \cap (\Phi_t \circ \gamma_{(\cdot)})^{-1}([a - 1/n, a])\right\}\right)$, then Corollary 5.4 gives that $n \cdot m_{\mathbb{L}_{\Phi_t^{-1}([a-1/n, a]) \cap \bar{\Gamma}_a(1)}}$ must converge to

$$\lambda_t g(\cdot, t) q = \lambda_t m_{a,t},$$

and therefore $\hat{m}_{a,t} = \lambda_t m_{a,t}$. \square

At the beginning of this section we observed that

$$e_t \# \gamma_a = \left(\int \varrho_t(z) \hat{m}_{a,t}(dz) \right)^{-1} \varrho_t \hat{m}_{a,t}.$$

Then Theorem 5.5 implies the next corollary.

Corollary 5.6. *The measure $(e_t) \# \gamma_a$ is absolute continuous with respect to the surface measure $m_{a,t}$.*

Let $\hat{h}_{a,t}(x) := \frac{d(e_t)_\# \gamma_a}{dm_{a,t}}(x)$ denote the density. Clearly $\hat{h}_{a,t}$ can be defined arbitrarily outside $e_t(G_a)$. Therefore for \mathcal{L}^1 -a.e. $t \in [0, 1]$, $(e_t)_\# \gamma_a = \hat{h}_{a,t} m_{a,t}$. We prefer to think of $\hat{h}_{a,t}$ as a function defined on G rather than on $e_t(G_a)$, hence define $h_{a,r} : G_a \rightarrow [0, \infty]$ by $h_{a,r}(\gamma) := \hat{h}_{a,r}(\gamma_r)$. So we have found a decomposition of ϱ_t :

$$\varrho_t(\gamma_t) = \left(\int \varrho_t(z) \hat{m}_{a,t}(dz) \right) \frac{1}{\lambda_t(\gamma_t)} h_{a,t}(\gamma),$$

where $a = \varphi(\gamma_0)$.

6. ESTIMATE IN CODIMENSION ONE

For $r \in [0, 1]$ and $a \in \mathbb{R}$, let $p_{a,r} := (e_t)_\# \gamma_a$. Consider $H \subset G$, γ -measurable with $\gamma(H) > 0$ and numbers $0 < R_0, L_0, R_1, L_1 < 1$ with $R_0 < R_1$ such that $R_t + L_t < 1$ for all $t \in [0, 1]$ where $R_t := (1-t)R_0 + tR_1$ and $L_t := (1-t)L_0 + tL_1$, then the following holds.

Lemma 6.1. *The curve*

$$(6.1) \quad t \mapsto \mu_t := \frac{1}{L_t \gamma(H)} \int_0^1 1_{(R_t, R_t + L_t) \times H}(e^{-1}(x)) p_{a,r}(dx) \mathcal{L}^1(dr) \in \mathcal{P}(M)$$

is a geodesic.

Proof. Observe that coupling each $\gamma_{R_s + \lambda L_s}$ with $\gamma_{R_t + \lambda L_t}$ for $\lambda \in [0, 1], \gamma \in H$ we obtain a d^2 -cyclically monotone coupling of μ_s with μ_t . The property then follows straightforwardly. \square

Hence, an optimal transport is achieved by not changing the “angular” parts and coupling radial parts according to optimal coupling on \mathbb{R} . We now prove that actually this is the also the only one.

Lemma 6.2. *For every $0 < R_0, L_0, R_1, L_1 < 1$ with $R_0 < R_1$ the geodesic $[0, 1] \ni t \mapsto \mu_t$ defined in (6.1) is the only geodesic between μ_0 and μ_1 .*

Proof. We show that $[0, 1] \ni t \mapsto \mu_t$ is the restriction of a longer geodesic, and, due to the non-branching property of the Wasserstein space inherited from (X, d, m) , this implies the claim.

Given $0 < R_0, L_0, R_1, L_1 < 1$ with $R_0 < R_1$ for $\varepsilon > 0$ consider

$$s := \frac{\varepsilon}{R_1 - R_0 + \varepsilon}, \quad L_\varepsilon := \frac{L_0 - sL_1}{1 - s}.$$

Clearly ε is chosen such that $L_\varepsilon > 0$. Then we can write the interval $[R_0, R_0 + L_0]$ as an inner point of a convex combination between interval:

$$[R_0 - \varepsilon, R_0 - \varepsilon + L_\varepsilon](1 - s) + [R_1, R_1 + L_1]s = [R_0, R_0 + L_0].$$

Replacing R_0 with $R_0 - \varepsilon$ and L_0 with L_ε in (6.1) and denote with $\hat{\mu}_t$ the new geodesic. It is straightforward to check that $\hat{\mu}_s = \mu_0$. The claim follows reasoning in the same way for R_1 and L_1 . \square

Observe that for each $t \in [0, 1]$ the density $\varrho_t(x)$ of μ_t w.r.t. m is given by

$$(6.2) \quad \varrho_t(\gamma_r) = \begin{cases} \frac{1}{L_t \gamma(H)} h_r(\gamma), & (r, \gamma) \in [R_t, R_t + L_t] \times H, \\ 0, & \text{otherwise.} \end{cases}$$

The following regularity result for densities holds true.

Lemma 6.3. *For γ_a -a.e. $\gamma \in G_a$, the function $r \mapsto h_{a,r}^{-1/N}(\gamma)$ is semi-concave on $(0, 1)$ and satisfies in distributional sense*

$$\partial_r^2 h_{a,r}^{-1/N}(\gamma) \leq -L^2(\gamma) \frac{K}{N} h_{a,r}^{-1/N}(\gamma),$$

where $L(\gamma) = d(\gamma_0, \gamma_1)$.

Proof. Recall that $\text{CD}_{loc}(K, N)$ implies $\text{CD}^*(K, N)$. Consider the geodesic μ_t defined in (6.1) with $L_0 = L_1$. From Lemma 6.2 we can apply the definition of $\text{CD}^*(K, N)$ to μ_t and get

$$(6.3) \quad h_{a,s}^{-1/N}(\gamma) \geq \frac{\sin((t-s)L(\gamma)\sqrt{K/N})}{\sin((t-r)L(\gamma)\sqrt{K/N})} h_{a,r}^{-1/N}(\gamma) + \frac{\sin((s-r)L(\gamma)\sqrt{K/N})}{\sin((t-r)L(\gamma)\sqrt{K/N})} h_{a,t}^{-1/N}(\gamma),$$

for all $0 < r < s < t < 1$ and γ_a -a.e. $\gamma \in G$. The claim is equivalent to (6.3). \square

Fix an open set $H \subset G$ and $[a, b] \subset [0, R]$ such that the curvature dimension condition $\text{CD}(K, N)$ holds true for all measures μ_0, μ_1 supported in $e([a, b] \times \bar{H})$. For each $R_0, R_1 \in (a, b)$ choose L_0, L_1 such that $R_0 + L_0, R_1 + L_1 \leq b$ and define $(\mu_t)_{t \in [0, 1]}$ as before in (6.1). Moreover we have to consider the following map

$$\begin{aligned} \Phi : \mathcal{G}(M) \times [0, 1] &\rightarrow \mathcal{G}(M) \\ (\gamma, s) &\mapsto t \mapsto \eta_t = \gamma_{(1-t)(R_0+sL_0)+t(R_1+sL_1)} \end{aligned}$$

Consider

$$\tilde{\gamma}_a := \Phi_{\#} \left(\frac{1}{\gamma_a(H)} \gamma_{a \perp H} \otimes \mathcal{L}^1_{[0, 1]} \right),$$

then $\mu_t = e_{t\#} \tilde{\gamma}_a$.

Theorem 6.4. *For γ_a -a.e. $\gamma \in H$ and for sufficiently close $0 \leq R_0 < R_1 \leq 1$ the following holds true:*

$$(6.4) \quad h_{a, R_{1/2}}^{-\frac{1}{N-1}}(\gamma) \geq \sigma_{K, N-1}^{(1/2)}((R_1 - R_0)L(\gamma)) \left\{ h_{a, R_0}^{-\frac{1}{N-1}}(\gamma) + h_{a, R_1}^{-\frac{1}{N-1}}(\gamma) \right\}.$$

Proof. Consider $0 \leq R_0 < R_1 < 1$, the measures μ_0, μ_1 , the corresponding measure on the space of geodesics $\tilde{\gamma}_a$ and recall that $\mu_t = \varrho_t m$.

Step 1. Condition $\text{CD}_{\text{loc}}(K, N)$ for $t = 1/2$ and the assumptions on R_0, L_0 and R_1, L_1 imply that for $\tilde{\gamma}_a$ -a.e. $\eta \in \mathcal{G}(M)$

$$\varrho_{1/2}^{-1/N}(\eta_{1/2}) \geq \tau_{K, N}^{(1/2)}(d(\eta_0, \eta_1)) \left\{ \varrho_0^{-1/N}(\eta_0) + \varrho_1^{-1/N}(\eta_1) \right\},$$

that can be formulated also in the following way: for \mathcal{L}^1 -a.e. $s \in [0, 1]$ and γ_a -a.e. $\gamma \in H$

$$\varrho_{1/2}^{-1/N}(\gamma_{R_{1/2}+sL_{1/2}}) \geq \tau_{K, N}^{(1/2)}((R_1 - R_0 + s|L_1 - L_0|)L(\gamma)) \left\{ \varrho_0^{-1/N}(\gamma_{R_0+sL_0}) + \varrho_1^{-1/N}(\gamma_{R_1+sL_1}) \right\}.$$

Then using (6.2) and the continuity of $r \mapsto h_r(\gamma)$ (Proposition 6.3), letting $s \searrow 0$, it follows that

$$(6.5) \quad (L_0 + L_1)^{1/N} h_{a, R_{1/2}}^{-1/N}(\gamma) \geq \sigma_{K, N-1}^{(1/2)}((R_1 - R_0)L(\gamma))^{\frac{N-1}{N}} \left\{ L_0^{1/N} h_{a, R_0}^{-1/N}(\gamma) + L_1^{1/N} h_{a, R_1}^{-1/N}(\gamma) \right\}$$

for all $R_0 < R_1 \in (a, b)$, all sufficiently small L_0, L_1 and γ_a -a.e. $\gamma \in H$, with exceptional set depending on R_0, R_1, L_0, L_1 .

Step 2. Note that all the involved quantities in (6.5) are continuous w.r.t. R_0, R_1, L_0, L_1 , therefore there exists a common exceptional set $H' \subset H$ of zero γ_a -measure such that (6.5) holds true for all for all $R_0 < R_1 \in (a, b)$, all sufficiently small L_0, L_1 and all $\gamma \in H \setminus H'$. Then for fixed $R_0 < R_1 \in (a, b)$ and fixed $\gamma \in H \setminus H'$, varying L_0, L_1 in (6.5) yields

$$h_{a, R_{1/2}}^{-\frac{1}{N-1}}(\gamma) \geq \sigma_{K, N-1}^{(1/2)}((R_1 - R_0)L(\gamma)) \left\{ h_{a, R_0}^{-\frac{1}{N-1}}(\gamma) + h_{a, R_1}^{-\frac{1}{N-1}}(\gamma) \right\}.$$

Indeed the optimal choice is

$$L_0 = L \frac{h_{a, R_0}^{-1/(N-1)}(\gamma)}{h_{a, R_0}^{-1/(N-1)}(\gamma) + h_{a, R_1}^{-1/(N-1)}(\gamma)}, \quad L_1 = L \frac{h_{a, R_1}^{-1/(N-1)}(\gamma)}{h_{a, R_0}^{-1/(N-1)}(\gamma) + h_{a, R_1}^{-1/(N-1)}(\gamma)}$$

for sufficiently small $L > 0$.

Repeating the same arguments with $[R_0 - L_0, R_0]$ and $[R_1 - L_1, R_1]$ we obtain also the case $R_1 = 1$. \square

7. GLOBAL ESTIMATE IN CODIMENSION ONE

From Theorem 6.4 we have that for γ_a -a.e. fixed $\gamma \in G_a \setminus H'$: for every $0 \leq R_0 < 1$ there exists $\varepsilon > 0$ such that for all $R_0 < R_1 < R_0 + \varepsilon \leq 1$ it holds

$$h_{a, R_{1/2}}^{-\frac{1}{N-1}}(\gamma) \geq \sigma_{K, N-1}^{(1/2)}((R_1 - R_0)L(\gamma)) \left\{ h_{a, R_0}^{-\frac{1}{N-1}}(\gamma) + h_{a, R_1}^{-\frac{1}{N-1}}(\gamma) \right\}.$$

We prove that mid-points inequality is equivalent to the complete inequality.

Lemma 7.1 (Midpoints). *Inequality (6.4) holds true if and only if*

$$(7.1) \quad h_{a,R_t}^{-\frac{1}{N-1}}(\gamma) \geq \sigma_{K,N-1}^{(1-t)}((R_1 - R_0)L(\gamma))h_{a,R_0}^{-\frac{1}{N-1}}(\gamma) + \sigma_{K,N-1}^{(t)}((R_1 - R_0)L(\gamma))h_{a,R_1}^{-\frac{1}{N-1}}(\gamma)$$

for all $t \in [0, 1]$.

Proof. We only consider the case $K > 0$. The general case requires analogous calculations. In order to make notations simpler we remove the subscripts from h_{a,R_t} .

Fix $0 \leq R_0 \leq R_1 \leq 1$, put $\theta := (R_1 - R_0)L(\gamma)$ and $h(s) := h_s(\gamma) = h(\gamma(s))$.

Step 1. For every $k \in \mathbb{N}$ we have

$$\begin{aligned} h^{-\frac{1}{N-1}}(R_0 + l2^{-k}\theta) &\geq \sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta)h^{-\frac{1}{N-1}}(R_0 + (l-1)2^{-k}\theta) \\ &\quad + \sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta)h^{-\frac{1}{N-1}}(R_0 + (l+1)2^{-k}\theta), \end{aligned}$$

for every odd $l = 0, \dots, 2^k$.

Step 2. We perform an induction argument on k : suppose that inequality (7.1) is satisfied for all $t = l2^{-k+1} \in [0, 1]$ with l odd, then (7.1) is verified by every $t = l2^{-k} \in [0, 1]$ with l odd:

$$\begin{aligned} h^{-\frac{1}{N-1}}(R_0 + l2^{-k}\theta) &\geq \sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta)h^{-\frac{1}{N-1}}(R_0 + (l-1)2^{-k}\theta) \\ &\quad + \sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta)h^{-\frac{1}{N-1}}(R_0 + (l+1)2^{-k}\theta) \\ &\geq \sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta) \left[h^{-\frac{1}{N-1}}(R_0)\sigma_{K,N-1}^{(1-(l-1)2^{-k})}(\theta) + h^{-\frac{1}{N-1}}(R_1)\sigma_{K,N-1}^{((l-1)2^{-k})}(\theta) \right] \\ &\quad + \sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta) \left[h^{-\frac{1}{N-1}}(R_0)\sigma_{K,N-1}^{(1-(l+1)2^{-k})}(\theta) + h^{-\frac{1}{N-1}}(R_1)\sigma_{K,N-1}^{((l+1)2^{-k})}(\theta) \right]. \end{aligned}$$

Following the calculation of the proof of Proposition 2.10 of [4], one obtain that

$$h^{-\frac{1}{N-1}}(R_0 + l2^{-k}\theta) \geq \sigma_{K,N-1}^{(1-l2^{-k})}(\theta)h^{-\frac{1}{N-1}}(R_0) + \sigma_{K,N-1}^{(l2^{-k})}(\theta)h^{-\frac{1}{N-1}}(R_1).$$

The claim is easily proved by the continuity of h and σ . □

We prove that (7.1) satisfies a local-to-global property.

Theorem 7.2 (Local to Global). *Suppose that for every $r \in [0, 1]$ there exists $\varepsilon > 0$ such that whenever $0 \leq r - \varepsilon \leq R_0 < R_1 \leq r + \varepsilon \leq 1$ then (7.1) holds true for all $t \in [0, 1]$. Then (7.1) holds true for all $0 \leq R_0 < R_1 \leq 1$ and all $t \in [0, 1]$.*

Proof. We only consider the case $K > 0$. The general case requires analogous calculations. Fix $0 \leq R_0 < R_1 \leq 1$, $\theta := (R_1 - R_0)L(\gamma)$ and $h(s) := h_s(\gamma) = h(\gamma(s))$.

Step 1. According to our assumption, every point $r \in [0, 1]$ has a neighborhood $(r - \varepsilon(r), r + \varepsilon(r))$ such that if R_0 and R_1 belong to that neighborhood then (7.1) is verified. By compactness there exist x_1, \dots, x_n such that the family $\{B_{\varepsilon(x_i)/2}(x_i)\}_{i=1,\dots,n}$ is a covering of $[0, 1]$. Let $\lambda := \min\{\varepsilon(x_i)/2 : i = 1, \dots, n\}$. Possibly taking a lower value for λ , we assume that $\lambda = 2^{-k}\theta$. Hence we have

$$\begin{aligned} h^{-\frac{1}{N-1}}(R_0 + \frac{1}{2}\theta) &\geq \sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta)h^{-\frac{1}{N-1}}(R_0 + \frac{1}{2}\theta - 2^{-k}\theta) \\ &\quad + \sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta)h^{-\frac{1}{N-1}}(R_0 + \frac{1}{2}\theta + 2^{-k}\theta). \end{aligned}$$

Step 2. We iterate the above inequality:

$$\begin{aligned}
h^{-\frac{1}{N-1}}(R_0 + \frac{1}{2}\theta) &\geq \sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta)h^{-\frac{1}{N-1}}(R_0 + \frac{1}{2}\theta - 2^{-k}\theta) \\
&\quad + \sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta)h^{-\frac{1}{N-1}}(R_0 + \frac{1}{2}\theta + 2^{-k}\theta) \\
&\geq \sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta) \left[\sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta)h^{-\frac{1}{N-1}}(R_0 + \frac{1}{2}\theta - 2^{-k+1}\theta) \right. \\
&\quad \left. + \sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta)h^{-\frac{1}{N-1}}(R_0 + \frac{1}{2}\theta) \right] \\
&\quad + \sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta) \left[\sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta)h^{-\frac{1}{N-1}}(R_0 + \frac{1}{2}\theta) \right. \\
&\quad \left. + \sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta)h^{-\frac{1}{N-1}}(R_0 + \frac{1}{2}\theta + 2^{-k+1}\theta) \right] \\
&\geq \sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta)^2 h^{-\frac{1}{N-1}}(R_0 + \frac{1}{2}\theta - 2^{-k+1}\theta) \\
&\quad + \sigma_{K,N-1}^{(1/2)}(2^{-k+1}\theta)^2 h^{-\frac{1}{N-1}}(R_0 + \frac{1}{2}\theta + 2^{-k+1}\theta).
\end{aligned}$$

Observing that $\sigma_{K,N-1}^{(1/2)}(\alpha)^2 \geq \sigma_{K,N-1}^{(1/2)}(2\alpha)$, it is fairly easy to obtain:

$$\begin{aligned}
h^{-\frac{1}{N-1}}(R_0 + \frac{1}{2}\theta) &\geq \sigma_{K,N-1}^{(1/2)}(2^{-k+i+1}\theta)h^{-\frac{1}{N-1}}(R_0 + \frac{1}{2}\theta - 2^{-k+i}\theta) \\
&\quad + \sigma_{K,N-1}^{(1/2)}(2^{-k+i+1}\theta)h^{-\frac{1}{N-1}}(R_0 + \frac{1}{2}\theta + 2^{-k+i}\theta),
\end{aligned}$$

for every $i = 0, \dots, k$. For $i = k - 1$ Lemma 7.1 implies the claim. \square

8. ESTIMATE IN DIMENSION ONE AND CONCLUSION

Recall the definition of $\lambda_t(x)$

$$\lambda_t(x) := \lim_{n \rightarrow \infty} n \cdot \mathcal{L}^1 \left(\left\{ \tau \in [0, 1] \cap (\Phi_t \circ \gamma_{(\cdot)})^{-1}([\varphi(\gamma_0) - 1/n, \varphi(\gamma_0)]) \right\} \right),$$

where γ is the unique element of the set $G \cap e_t^{-1}(x)$.

Without any other assumption on the space, in order to prove concavity in time of $\lambda_t(\gamma_t)$, we have to make an assumption on the shape of the chosen optimal dynamical transference plan γ .

Proposition 8.1. *Assume the following: for γ -a.e. $\gamma \in G$, $d^2(\gamma_0, \gamma_1)$ depends only on $\varphi(\gamma_0)$, i.e.*

$$d(\gamma_0, \gamma_1) = d(\gamma'_0, \gamma'_1),$$

whenever $\varphi(\gamma_0) = \varphi(\gamma'_0)$. Then for γ -a.e. $\gamma \in G$ the following holds true

$$\lambda_t(\gamma_t) = (1-t)\lambda_0(\gamma_0) + t\lambda_1(\gamma_1),$$

for every $t \in [0, 1]$.

Proof. Step 1. From the hypothesis follow that $d^2(\gamma_0, \gamma_1)$ is constant for γ_0 varying over a level set of φ . It follows that $\{\gamma_t : \varphi(\gamma_0) = a\}$ is a level set of φ_t indeed

$$\varphi_t(\gamma_t) = (1-t)\varphi(\gamma_0) + t\varphi_1(\gamma_1), \quad \gamma - a.e. \gamma \in G,$$

where $\varphi_1 = -\varphi^d$. We can therefore consider the disintegration induced by the level sets of φ_t

$$m_{\perp_{e_t}(G)} = \int_{\varphi_t(\mu_t)} \check{m}_{t,b} \mathcal{L}^1(db).$$

Since

$$\lim_{\tau \rightarrow 0} \frac{\varphi_t(\gamma_t) - \varphi_t(\gamma_{t+\tau})}{\tau} = \frac{d^2(\gamma_t, \gamma_1)}{(1-t)^2} = d^2(\gamma_0, \gamma_1),$$

repeating the very same proof of Theorem 5.5 with φ_t instead of Φ_t , we obtain

$$\check{m}_{t,\xi_t(a)} = \frac{1}{d^2(\gamma_0, \gamma_1)} m_{a,t},$$

where $\xi_t(a)$ is the unique element of $\varphi_t(\mu_t)$ such that $\varphi_t^{-1}(\xi_t(a)) = \{\gamma_t : \varphi_{\gamma_0} = a\}$.

Step 2. Since they represent the same geodesic, the two measures

$$h_{a,t}m_{a,t}, \quad \frac{\varrho_t}{\int \varrho_t \tilde{m}_{t,\xi_t(a)}} \tilde{m}_{t,\xi_t(a)}$$

coincides. Now observe that the map ξ_t pushes the quotient measure $\varphi_{\#}\mu_0$ to the quotient measure $\varphi_t\#\mu_t$ in a monotone way. From one dimensional theory, it holds that $\xi_t = (1-t)Id + t\xi_1$ and

$$|\partial_a \xi_t|(a) \int \varrho_t \tilde{m}_{t,\xi_t(a)} = \int \varrho_0 \tilde{m}_{0,a} = \int \varrho_0 \hat{m}_{a,0}.$$

Hence

$$\varrho_t = d^2(\gamma_0, \gamma_1) h_{a,t} \int \varrho_t \tilde{m}_{t,\xi_t(a)} = h_{a,t} \int \varrho_0 \hat{m}_{a,0} \frac{d^2(\gamma_0, \gamma_1)}{|\partial_a \xi_t|(a)},$$

and therefore, up to constant, since $\int \varrho_t \hat{m}_{a,t}$ is constant in t ,

$$\lambda_t(\gamma_t) = |\partial_a \xi_t|(a),$$

and the concavity is therefore proved. \square

Instead of making assumption on γ we can make an assumption on (X, d, m) . We will ask the metric measure space (X, d, m) to be *infinitesimally strictly convex* in the sense of Definition 2.12. This will implies better properties for λ_t .

Define the restriction map as follows. For any $t \in (0, 1)$ let $restr_{[t,1]} : \mathcal{G}(X) \rightarrow \mathcal{G}(X)$ be defined as follows $restr_{[t,1]}(\gamma)_s = \gamma_{(1-s)t+s}$. Denote by $\gamma_{[t,1]}$ the measure $restr_{[t,1]}\#\gamma$.

Lemma 8.2. *For all $t \in [0, 1)$ the measure $\gamma_{[t,1]}$ represents $\nabla(1-t)\varphi_t$.*

The notion of test plans representing gradients has been introduced in Definition 2.10.

Proof. First observe that $\varphi_t \in S^2(e_t(G), d, m)$. Indeed from Proposition 2.8, since φ_t is a Kantorovich potential for (μ_t, μ_1) , it follows that

$$|D\varphi_t|_w(\gamma_t) = \frac{d(\gamma_t, \gamma_1)}{1-t} = d(\gamma_0, \gamma_1), \quad \text{for } \gamma - a.e. \gamma,$$

and therefore $|D\varphi_t|_w \in L^2(e_t(G), m)$. We know that $\gamma_{[t,1]}$ is the optimal dynamical transference plan between μ_t and μ_1 and $(1-t)\varphi_t$ is the Kantorovich potential for the d^2 cost, hence Proposition 2.8 implies that

$$\lim_{t \downarrow 0} \int \frac{\varphi_t(\gamma_\tau) - \varphi_t(\gamma_0)}{\tau} \gamma_{[t,1]}(d\gamma) = \int \frac{d^2(\gamma_0, \gamma_1)}{1-t} \gamma_{[t,1]}(d\gamma).$$

Since $\|\gamma_{[t,1]}\|_2^2 = \int d^2(\gamma_0, \gamma_1) \gamma_{[t,1]}(d\gamma)$,

$$(1-t) \lim_{t \downarrow 0} \int \frac{\varphi_t(\gamma_\tau) - \varphi_t(\gamma_0)}{\tau} \gamma_{[t,1]}(d\gamma) = \|\gamma_{[t,1]}\|_2^2$$

and the claim follows. \square

Using Theorem 2.11 we can now express $L_\gamma(t)$ in terms of a second order differentiation.

Proposition 8.3. *Let (X, d, m) be infinitesimally strictly convex. Then λ_t verifies the following identity: for every $t \in [0, 1)$*

$$\frac{1}{\lambda_t(\gamma_t)} = -D\Phi_t(\nabla\varphi_t)(\gamma_t), \quad \gamma - a.e. \gamma,$$

where the exceptional set depends on t .

Proof. Step 1. Since the support of γ , $G \subset \mathcal{G}(M)$, is compact, φ_t is Lipschitz on $e_t(G)$ with finite Lipschitz constant. Moreover

$$d^2(\gamma_t, \gamma_1) = (1-t) \sup_{y \in M} \{\varphi_t(\gamma_t) + \varphi^d(y)\},$$

hence Φ_t is Lipschitz as well, and from the compactness of $e_t(G)$, it belongs to $S^2(e_t(G), d, m)$.

Step 2. Since (X, d, m) is infinitesimally Hilbertian, and $\gamma_{[t,1]}$ represents $\nabla(1-t)\varphi_t$, from Theorem 2.11 it follows that

$$\begin{aligned} \lim_{\tau \downarrow 0} \int_{\text{restr}_{[t,1]}(G)} \frac{\Phi_t(\gamma_\tau) - \Phi_t(\gamma_0)}{\tau} \gamma_{[t,1]}(d\gamma) &= (1-t) \int D\Phi_t(\nabla\varphi_t)(x) \mu_t(dx) \\ &= (1-t) \int_{\text{restr}_{[t,1]}(G)} D\Phi_t(\nabla\varphi_t)(\gamma_0) \gamma_{[t,1]}(d\gamma). \end{aligned}$$

Since the previous identity holds true even if we restrict to a subset of $\text{restr}_{[t,1]}(G)$, it follows that it holds point-wise: for $\gamma_{[t,1]}$ -a.e. γ

$$\lim_{\tau \downarrow 0} \frac{\Phi_t(\gamma_\tau) - \Phi_t(\gamma_0)}{\tau} = (1-t) D\Phi_t(\nabla\varphi_t)(\gamma_0).$$

So fix $\hat{\gamma}$ in the support of $\gamma_{[t,1]}$ such that the limit exists and consider γ in the support of γ such that $\hat{\gamma}_\tau = \gamma_{(1-\tau)t+\tau}$, then we have

$$\frac{\Phi_t(\hat{\gamma}_\tau) - \Phi_t(\hat{\gamma}_0)}{\tau} = \frac{\Phi_t(\gamma_{(1-\tau)t+\tau}) - \Phi_t(\gamma_t)}{\tau} = \frac{\Phi_t(\gamma_{(1-\tau)t+\tau}) - \Phi_t(\gamma_t)}{\tau(1-t)}(1-t),$$

and therefore the claim follows. \square

Under the infinitesimally Hilbertian assumption, we have therefore the following decomposition:

$$\frac{1}{c(\varphi(\gamma_0))} \varrho_t(\gamma_t) = -D\Phi_t(\nabla\varphi_t)(\gamma_t) h_t(\gamma_t),$$

where $c(a) = \int \varrho_t(z) \hat{m}_{a,t}(dz)$ is independent of t and h verifies $\text{CD}^*(K, N-1)$.

Using the results proved so far, we can now state the following.

Theorem 8.4. *Assume that, for γ -a.e. $\gamma \in G$, $d^2(\gamma_0, \gamma_1)$ depends only on $\varphi(\gamma_0)$. Then the following holds true*

$$\varrho_t(\gamma_t)^{-1/N} \geq \varrho_0(\gamma_0)^{-1/N} \tau_{K,N}^{(1-t)}(d(\gamma_0, \gamma_1)) + \varrho_1(\gamma_1)^{-1/N} \tau_{K,N}^{(s)}(d(\gamma_0, \gamma_1)),$$

for every $t \in [0, 1]$ and for γ -a.e. $\gamma \in G$.

Proof. Since

$$\varrho_t(\gamma_t) = \left(\int \varrho_t(z) \hat{m}_{a,t}(dz) \right) \frac{h_{a,t}(\gamma)}{\lambda_t(\gamma_t)},$$

where the integral is constant in t , in order to prove the claim we can assume

$$\varrho_t(\gamma_t) = \frac{1}{\lambda_t(\gamma_t)} h_{a,t}(\gamma).$$

Then from Theorem 6.4 and Proposition 8.1

$$\begin{aligned} \varrho_t^{-1/N}(\gamma_t) &= \left(\frac{1}{\lambda_t(\gamma_t)} h_{a,t}(\gamma) \right)^{-1/N} \\ &= \left((1-t)\lambda_0(\gamma_0) + t\lambda_1(\gamma_1) \right)^{\frac{1}{N}} \left(h_{a,t}^{-1/(N-1)}(\gamma) \right)^{\frac{N-1}{N}} \\ &\geq \left((1-t)\lambda_0(\gamma_0) \right)^{1/N} \left(\sigma_{K,N-1}^{(1-t)}(d(\gamma_0, \gamma_1)) h_{a,0}^{-\frac{1}{N-1}}(\gamma) \right)^{\frac{N-1}{N}} \\ &\quad + \left(t\lambda_1(\gamma_1) \right)^{1/N} \left(\sigma_{K,N-1}^{(t)}(d(\gamma_0, \gamma_1)) h_{a,1}^{-\frac{1}{N-1}}(\gamma) \right)^{\frac{N-1}{N}} \\ &= \varrho_0^{-1/N}(\gamma_0) \tau_{K,N}^{(1-t)}(d(\gamma_0, \gamma_1)) + \varrho_1^{-1/N}(\gamma_1) \tau_{K,N}^{(t)}(d(\gamma_0, \gamma_1)). \end{aligned}$$

The claim follows. \square

REFERENCES

- [1] L. Ambrosio and N. Gigli. A user's guide to optimal transport. preprint, 2011.
- [2] L. Ambrosio, N. Gigli, and G. Savaré. Calculus and heat flow in metric measure spaces and application to spaces with Ricci curvature bounded from below. Preprint, arXiv:1106.2090.
- [3] L. Ambrosio, M. Miranda Jr., and D. Pallara. Special functions of bounded variation in doubling metric measure spaces. *Calculus of variations: topics from the mathematical heritage of E. De Giorgi*, pages 1–45, 2004.
- [4] K. Bacher and K.T. Sturm. Localization and tensorization properties of the Curvature-Dimension condition for metric measure spaces. *J. Funct. Anal.*, 259(1):28–56, 2010.
- [5] D. Bakry and Z. Qian. Some new results on eigenvectors via dimension, diameter, and Ricci curvature. *Adv. Math.*, 155:98–153, 2000.
- [6] S. Bianchini and L. Caravenna. On the extremality, uniqueness and optimality of transference plans. *Bull. Inst. Math. Acad. Sin.(N.S.)*, 4(4):353–454, 2009.
- [7] S. Bianchini and F. Cavalletti. The Monge problem for distance cost in geodesic spaces. Accepted on *Comm. Math. Phys.*, 2010.
- [8] F. Cavalletti and K.-T. Sturm. Local curvature-dimension condition implies measure-contraction property. *J. Funct. Anal.*, 262:5110 – 5127, 2012.
- [9] D. H. Fremlin. *Measure Theory*, volume 4. Torres Fremlin, 2002.
- [10] N. Gigli. Optimal maps in non branching spaces with ricci curvature bounded from below. preprint,, 2011.
- [11] N. Gigli. On the differential structure of metric measure spaces and applications. preprint, arXiv:1205.6622, 2012.
- [12] J. Lott and C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math.*, 169(3):903–991, 2009.
- [13] K.T. Sturm. On the geometry of metric measure spaces.I. *Acta Math.*, 196(1):65–131, 2006.
- [14] K.T. Sturm. On the geometry of metric measure spaces.II. *Acta Math.*, 196(1):133–177, 2006.
- [15] C. Villani. *Optimal transport, old and new*. Springer, 2008.

HAUSDORFF CENTER FOR MATHEMATICS, ENDENICHER ALLEE 62, D-53115 BONN (GERMANY)
 E-mail address: `fabio.cavalletti@hcm.uni-bonn.de`